

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 23:11:27 ; Search time 230 Seconds
(without alignments)
1165.655 Million cell updates/sec

Title: US-10-763-042-9

Perfect score: 1962

Sequence: 1 MTFRSLQEMRSPHRVHVA.....SAFQAFALCSFFETRIACE 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1962	100.0	380	Q9SQU1_ARATH	Q9squ1 arabidopsis
2	1442.5	73.5	389	Q6NQP1_ARATH	Q6npq1 arabidopsis
3	1414.5	72.1	380	Q6YXN7_ARATH	Q6ynx7 arabidopsis
4	1149	58.6	406	Q8VY21_ARATH	Q8vy21 arabidopsis
5	1147.5	58.5	407	Q82257_ARATH	Q82257 arabidopsis
6	1139	58.1	406	Q8H485_ORYSA	Q8h485 oryza sativ
7	1015	51.7	451	Q69U54_ORYSA	Q69u54 oryza sativ
8	1013.5	51.7	445	Q75HX5_ORYSA	Q75hx5 oryza sativ
9	1013.5	51.7	455	Q9ZP59_ARATH	Q9zps9 arabidopsis
10	999.5	50.9	445	Q9FRH7_ARATH	Q9frh7 arabidopsis
11	992	50.6	448	Q8LTA9_ORYSA	Q8lta9 oryza sativ
12	989	50.4	462	Q7XSU4_ORYSA	Q7xsu4 oryza sativ
13	989	50.4	462	Q9ST78_ORYSA	Q9st78 oryza sativ
14	969	49.4	440	Q53FP5_ORYSA	Q53fp5 oryza sativ
15	966.5	49.3	394	Q8GVE5_ARATH	Q8gve5 arabidopsis
16	957.5	48.8	394	Q7Y213_ARATH	Q7y213 arabidopsis
17	950.5	48.4	429	Q94483_ARATH	Q94483 arabidopsis
18	937	47.8	428	Q9SX13_9ARAE	Q9sx13 lemma pauci
19	933.5	47.6	411	Q9MJH9_CICAR	Q9mjh9 cicor ariet
20	929.5	47.4	415	Q84JM8_ARATH	Q84jm8 arabidopsis
21	929	47.3	388	Q9C6B4_ARATH	Q9c6b4 arabidopsis
22	927.5	47.3	413	Q84U70_ARALY	Q84u70 arabidopsis
23	927.5	47.3	415	Q84UG3_ARATH	Q84ug3 arabidopsis
24	925.5	47.2	415	Q84UG2_ARATH	Q84ug2 arabidopsis
25	925.5	47.2	428	Q6Z2G9_ORYSA	Q6z2g9 oryza sativ
26	919.5	46.9	386	Q9ZFW1_ARATH	Q9zfw1 arabidopsis
27	843	43.0	379	Q93VI8_ARATH	Q93vi8 arabidopsis
28	832	42.4	356	Q5QM27_ORYSA	Q5qm27 oryza sativ
29	816.5	41.6	579	Q9MAG9_ARATH	Q9mag9 arabidopsis
30	810.5	41.3	372	Q68Y48_ORYSA	Q68y48 oryza sativ
31	800.5	40.8	352	Q68BY7_ORYSA	Q68by7 oryza sativ

32	787.5	40.1	415	2	Q9XIP9_ARATH	Q9xip9 arabidopsis
33	723.5	36.9	317	2	Q53PP4_ORYSA	Q53pp4 oryza sativ
34	723	36.9	368	2	Q94DT9_ORYSA	Q94dt9 oryza sativ
35	637.5	32.5	267	2	Q8HOW5_ARATH	Q8how5 arabidopsis
36	596.5	30.4	219	2	Q8W220_PYRGO	Q8w220 pyrus commu
37	565.5	28.8	326	2	Q622H0_ORYSA	Q622h0 oryza sativ
38	456	23.2	486	2	Q5XFX6_RAT	Q5xfx6 rattus norv
39	455.5	23.2	438	2	Q5FW93_MOUSE	Q5fw93 mus musculu
40	450	22.9	540	2	Q4S4Q7_TETNG	Q4s4q7 tetraodon n
41	441.5	22.5	269	2	Q8C5R7_MOUSE	Q8c5r7 mus musculu
42	441.5	22.5	562	2	Q8C5R9_MOUSE	Q8c5r9 mus musculu
43	441.5	22.5	564	1	TULP2_MOUSE	P46686 mus musculu
44	440.5	22.5	506	2	Q6DBE2_XENLA	Q6dee2 xenopus lae
45	437.5	22.3	420	2	Q9H4D2_HUMAN	Q9h4d2 homo sapien

ALIGNMENTS

RESULT 1

Q9SQU1_ARATH PRELIMINARY; PRT; 380 AA.
AC Q9SQU1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE F24PI7.15 protein (Tubby-like protein TULP9).
GN Name=F24PI7.15; Synonym=TULP9;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15064372; DOI=10.1104/pp.103.037820;
RA Lai C.P., Lee C.L., Chen P.H., Wu S.H., Yang C.C., Shaw J.P.;
RT "Molecular analyses of the Arabidopsis TUBBY-like protein gene family.";
RL Plant Physiol. 134:1586-1597(2004).
DR EMBL; AC011623; AAF08576.1; -; Genomic_DNA.
DR EMBL; AF487270; AAQ06243.1; -; mRNA.
DR HSSP; P50586; 117E.
DR GO; GO:0045449; P:regulation of transcription; TAS.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB 1; UNKNOWN 1.
SQ SEQUENCE 380 AA; 42310 MW; EDB57859F8B5F51D CRC64;

Query Match 100.0%; Score 1962; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.3e-165;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MTFRSLQEMRSPHRVHVAHAASTANSSDPFSWSELPEELLREILIRVETVCGDWPSRR	60
Db	1	MTFRSLQEMRSPHRVHVAHAASTANSSDPFSWSELPEELLREILIRVETVCGDWPSRR	60
QY	61	NNVACAGVCRSWRLTKEIVAVPEPSSKLTFFISLQSGPRSLVQCFFKRNRNTOSYHL	120
Db	61	NNVACAGVCRSWRLTKEIVAVPEPSSKLTFFISLQSGPRSLVQCFFKRNRNTOSYHL	120
QY	121	YLGLTTLTDNGKFLLAASKLRATCTDYIIISLRSDDISKSNAYLGRMRSNFLGTKFTV	180
Db	121	YLGLTTLTDNGKFLLAASKLRATCTDYIIISLRSDDISKSNAYLGRMRSNFLGTKFTV	180

```
QY 181 FDGSOTGAAMOKRSRNFIVKSPVPOGSPYPIAHISYELNVLGSRGPRMRCIMDTIPM 240
Db 181 FDGSOTGAAMOKRSRNFIVKSPVPOGSPYPIAHISYELNVLGSRGPRMRCIMDTIPM 240
QY 241 SVESRGVAVASTISISSFSRSPVFRSHKPLRSNSASCDSGNNLGDPLVLNKAAPRW 300
Db 241 SVESRGVAVASTISISSFSRSPVFRSHKPLRSNSASCDSGNNLGDPLVLNKAAPRW 300
QY 301 HEQLRCWCLNFRGRTVAVSVKQFQVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
Db 301 HEQLRCWCLNFRGRTVAVSVKQFQVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
QY 361 SAFOAFAICLSSFETRIACE 380
Db 361 SAFOAFAICLSSFETRIACE 380

RESULT 2
Q6YNK7 ARATH PRELIMINARY; PRT; 389 AA.
AC Q6YNK7_ ARATH PRELIMINARY; PRT; 389 AA.
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
DE At5g18680.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chouk R., Chen H., Kim C.J., Shinn P., Ecker J.R.;
RA Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Kim C.J., Chen H., Chouk R., Shinn P., Ecker J.R.;
RA "Arabidopsis ORF clones.";
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT010768; AAR23738.1; -; mRNA.
DR EMBL; BT020601; AAW80874.1; -; mRNA.
DR GO; GO:0006512; P-ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB 1; UNKNOWN 1.
SQ SEQUENCE 389 AA; 43050 MW; 1270D94D2FB4E135 CRC64;

Query Match 73.5%; Score 1442.5; DB 2; Length 389;
Best Local Similarity 72.7%; Pred. No. 5.7e-119;
Matches 288; Conservative 32; Mismatches 53; Indels 23; Gaps 5;

QY 1 MTRSLLOEMRSPHRVH-----AAASTANSDDPFSWSELPEELLREILIRVETVDCG 54
Db 1 MTRSLILEMRSPHRVHDLAAAAADSTSVSQDYRWSEIPEELLREILIRVEAADGG 60
QY 55 DWPSRRNVACAGVCSWRILTKIIVAVPEFSSKLTPTISLKQSGPRDSLVCQFIKRNRN 114
Db 61 GWPSRRNVACAGVCGRWLLMNETVVVPEISKLTFPISLKQSGPRDSLVCQFIKRNRI 120
QY 115 TQSYHLYLGLTSLTDNGKFLLAASKLKRATCTDYIISLRSDDI SKRSNAYLGRMRSNFI 174
Db 121 TQSYHLYLGLTSLTDGKFLLAACKLHTTCTDYIISLRSDMSRRSQAYVGVKVSNEF 180
QY 175 GTKFTVFDGS---QTGAAMOKRSRNFIVKSPVPOGSPYPIAHISYELNVLGSRGPRM 231
Db 181 GTKFTVFDGNLSTGAALKRSRNPYPAKVSAPVPLGSPYVAHIITELNVLGSRGPRM 240
QY 232 RCIKMDTIPMSIVSRGVAVASTISISSFSRSPVFRSHKPLRSNSASCDSGNNLGDPL 291
Db 232 RCIKMDTIPMSIVSRGVAVASTISISSFSRSPVFRSHKPLRSNSASCDSGNNLGDPL 291
```

```
Db 241 QCLMDTIPTIMEPOGVASEPSEFFLLGLTRSTLRSQSOKPLRSSS-----SHLKETPL 293
QY 292 VLSNKAPEWHQRLRCWCLNFRGRTVAVSVKQFQVAV-SDCEAG-----QTSERILIQF 344
Db 294 VLSNKTPTWHQRLRCWCLNFRGRTVAVSVKQFQVAVAGASCGSGTGMSPERQSERIILQF 353
QY 345 GKVGKDMFTMDYGYPISAFOAFAICLSSFETRIACE 380
Db 354 GKVGKDMFTMDYGYPISAFOAFAICLSSFETRIACE 389

RESULT 3
Q6YNK7 ARATH PRELIMINARY; PRT; 380 AA.
AC Q6YNK7_ ARATH PRELIMINARY; PRT; 380 AA.
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Tubby-like protein 12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA PubMed=15064372; DOI=10.1104/pp.103.037820;
RA Lai C.P., Lee C.L., Chen P.H., Wu S.H., Yang C.C., Shaw J.F.;
RT "Molecular analyses of the Arabidopsis TUBBY-like protein gene family.";
RL Plant Physiol. 134:1586-1597 (2004).
DR EMBL; AV046922; AAL03978.1; -; mRNA.
DR GO; GO:0006512; P-ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB 1; UNKNOWN 1.
SQ SEQUENCE 380 AA; 41993 MW; EF52AA224CDA8F4D CRC64;

Query Match 72.1%; Score 1414.5; DB 2; Length 380;
Best Local Similarity 72.9%; Pred. No. 1.7e-116;
Matches 282; Conservative 31; Mismatches 51; Indels 23; Gaps 5;

QY 10 MRSRPHRVVH-----AAASTANSDDPFSWSELPEELLREILIRVETVDCGWPSPRVV 63
Db 1 MRSRPHRVVHDLAAAAADSTSVSQDYRWSEIPEELLREILIRVEAADGGWPSPRVV 60
QY 64 ACAGVCSWRILTKIIVAVPEFSSKLTPTISLKQSGPRDSLVCQFIKRNNTQSYHLYLG 123
Db 61 ACAGVCSWRILTKIIVAVPEFSSKLTPTISLKQSGPRDSLVCQFIKRNNTQSYHLYLG 120
QY 124 LTSLTDNGKFLLAASKLKRATCTDYIISLRSDDI SKRSNAYLGRMRSNFI 183
Db 121 LTSLTDGKFLLAACKLHTTCTDYIISLRSDMSRRSQAYVGVKVSNEF 180
QY 184 S---QTGAAMOKRSRNFIVKSPVPOGSPYPIAHISYELNVLGSRGPRMRCIMDTIPM 240
Db 181 NLLPSTGAALKRSRNPYPAKVSAPVPLGSPYVAHIITELNVLGSRGPRMRCIMDTIPT 240
QY 241 SVESRGVAVASTISISSFSRSPVFRSHKPLRSNSASCDSGNNLGDPLVLNKAAPRW 300
Db 241 STMEPOGVASEPSEFFLLGLTRSTLRSQSOKPLRSSS-----SHLKETPLVLNKTIPW 293
QY 301 HEQLRCWCLNFRGRTVAVSVKQFQVAV-SDCEAG-----QTSERILIQFGKVGDMFT 353
Db 294 HEQLRCWCLNFRGRTVAVSVKQFQVAVAGASCGSGTGMSPERQSERIILQFGKVGDMFT 353
QY 354 MDYGYPISAFOAFAICLSSFETRIACE 380
Db 354 MDYGYPISAFOAFAICLSSFETRIACE 380
```

RESULT 4
Q8VY21_ ARATH PRELIMINARY; PRT; 406 AA.
ID Q8VY21_ ARATH PRELIMINARY; PRT; 406 AA.
AC Q8VY21;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Tubby-like protein 3.
GN Name=At2g47900;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosoids II; Brassicales; Brassicaceae; Arabidopsi.s.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Becker J.R.,
RA Theologis A.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15064372; DOI=10.1104/pp.103.037820;
RA Lai C.P., Lee C.L., Chen P.H., Wu S.H., Yang C.C., Shaw J.F.;
RT "Molecular analyses of the Arabidopsis TUBBY-like protein gene
RT family";
RI Plant Physiol. 134:1586-1597(2004).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,
RA Davis R.W., Becker J.R., Theologis A.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY074273; AAL66970.1; -; mRNA.
DR EMBL; AY054774; AAK98802.1; -; mRNA.
DR EMBL; AY096604; AAM20254.1; -; mRNA.
DR HSSP; P50586; 117E.
DR GO; GO:0045449; P:regulation of transcription; TAS.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
SQ SEQUENCE 406 AA; 45311 MW; C9A8188C0A0D77D4 CRC64;
Query Match 58.6%; Score 1149; DB 2; Length 406;
Best Local Similarity 56.7%; Pred. No. 6, 1e-93;
Matches 233; Conservative 58; Mismatches 84; Indels 36; Gaps 9
Qy 1 MTFSLQLQEM-----RSRPHRVHAAASTANSDDPFWSSELPPELL 41
Dd : :: :: :
Db 1 MSPKSLIQDMRGELGSI SRKGDFVRFGVGSRSRQVVDTSVPVDAFKQSCWASMPPELL 60
Qy 42 REILIRVETVDGGDPSPERNVVACGVCSRWILTKEIVAVPEFSKLTFFLSLKQSGPR 101
Dd : :: :: :
Dd 61 RDVLMLRIEQSE-DTWPPSRKNVVSCAGVCNNRWREI VKEIVRVPESLSKLTFFLSLKQSGPR 119
Qy 102 DSLVQCFTKRNNTQSYYHLGLTTLSTL-DNGKFLLAASKLRATCTDYIIISLSDISK 160
Dd : :: :: :
Dd 120 GSLVQCYIMNRNSQTYLYLGLNQAAASNDGKFLLAARFRPTCTDYIIISLCDVSR 179
Qy 161 RSNYALGRMRSNF LGTKPTFDGSGQTG-AAKWQKSRRSNFI---KVSRPVQGSYPPIAH 216
Dd : :: :: :

180 GSNYYIGKLRNFI^{LG}TFTY^{VD}QA^{PT}NP^GTQ^{VT}TR^{SR}LL^{SL}KQ^{VS}PR^{IP}SG^{NP}YVA^{HI} 233

217 SYELNVLGSRGPR^{MR}CI^{MD}TI^{PM}SI^{VE}SRG^V-----VASTSI^{SS}FS^{SS}SR^{SS}SP^{VR}FRSHK 270

240 SYELNVLGSRGPR^{MR}QC^{MD}AI^{PA}SA^{VE}PG^{TA}PT^{QT}EL^{VH}NLD^{SP}FS⁻-F^{FR}--SK 295

271 PLRNSASCS^{DS}GNL^{GD}PP^{LV}LS^{NK}AP^{RW}HE^{OL}R^{CH}W^{GL}PH^{GT}V^{AS}VKN^{FOL}VAV^{-S} 329

296 SIRAES^{LP}SG^{SS}AA^{QK}EG^{LL}VL^{KN}KAP^{RW}HE^QQC^{WL}NP^{NG}RV^{TV}VAS^{VKN}FOL^{VAA}PE 355

330 DCEAG^QSE^{RI}IL^{FG}VK^{GD}MF^{TD}MY^{GY}PI^{SA}FA^{FC}AIL^{SS}SP^{ET}RIACE 380

356 NGPAG^{PE}HENV^{IL}Q^{FG}VK^{GD}VF^{TD}MY^{GY}PI^{SA}FA^{FC}IL^{SS}PD^{TT}KIACE 406

RESULT 5

082257 ARATH

ID 082257 ARATH PRELIMINARY; PRT; 407 AA.

AC 082257;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

DE Putative tubby protein.

GN Name=At2g47900;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI_TaxId=3702;

[1]

RP NUCLEOTIDE SEQUENCE.

RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,

RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[2]

RP NUCLEOTIDE SEQUENCE.

RA Town C.D., Kaul S.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

[3]

RP NUCLEOTIDE SEQUENCE.

RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.B.,

RA Barstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,

RA Carreia A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,

RA Fraser C.M., Venter J.C.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC005309; AAC63644.1; -; Genomic DNA.

DR EMBL; AC006072; AAM15124.1; -; Genomic_DNA.

DR PIR; H84920; H84920.

DR HSSP; P50586; 117E.

DR GO; GO:0006512; P:ubiquitin cycle; IEA.

DR InterPro; IPR001810; F-box.

DR InterPro; IPR000007; Tubby.

DR Pfam; PF00646; F-box; 1.

DR Pfam; PF01167; Tub; 1.

DR PRINTS; PR01573; SUPERTUBBY.

SQ SEQUENCE 407 AA; 45382 MW; 45AA55907F09DE86 CRC64;

Query Match 58.5%; Score 1147.5; DB 2; Length 407;

Best Local Similarity 56.6%; Pred. No. 8.3e-93;

Matches 233; Conservative 57; Mismatches 85; Indels 37; Gaps 9;

QY 1 MTFESLQLQEM-----SSRPHRVVHAAASTANSDDPFSWSLPELL 41

DB 1 MSPFKSLIQDMRGELGSLRGKGFVRFYGRSRSQVVDTSVPVDAFKQSCWASMPPELL 60

QY 42 REILIRIVETVDGGDPWPSRRNVACAGYCRSWRLITKEIVAVPEPSSKLTTPFISLKQSGPR 101

DB 61 RDVLMLRIEQSE-DTWPSRKNVSCAGYCRNWRREIVKELVIRVPELSKLTTPFISLKQGPGR 119

QY 102 DSLVQCPTKKNRNTQSYHLIYGL--TTLSDTNGKFLLAASKLKRAKTCITYIISLRSDDIS 159

DB 120 GSLVQCYIMRNRNOTYLYLGLNQAAASNDGDGFLLAAREFRPPTCTDYIISLNCDDVS 179

```
QY 160 KRSNAYLGRMRNFIQKTFVDSQTG-AAKMQKRSRNF1---KVSPRVPGSGYP1AH 215
Db 180 RGSNTYIGKLSNFIQKTFVDSQTG-AAKMQKRSRNF1---KVSPRVPGSGYP1AH 239
QY 216 ISYELNVLGSRGPRRMRIMDTIPMSIVSRGV-----VASTSISSPSRSPVPRSHS 269
Db 240 ISYELNVLGSRGPRRMRIMDTIPMSIVSRGV-----VASTSISSPSRSPVPRSHS 295
QY 270 KPLRSNSASCSGNNLGPPLVLSNKAPRWHEQLRCWCLNFGHGRVTVASVKNFQLVAV- 328
Db 296 KSIRAESLSPGSAQAQKGLLVLSNKAPRWHEQLRCWCLNFGHGRVTVASVKNFQLVAV- 355
QY 329 SDCQAGQTSERIILQFGKVKDMFTMDYGYPISAFOAFAICLSSFFETRIACE 380
Db 356 ENGPAGPEHNVILQFGKVKDMFTMDYGYPISAFOAFAICLSSFFETRIACE 407

RESULT 6
Q6H485 ORYSA
ID Q6H485_ORYSA PRELIMINARY; PRT; 406 AA.
AC Q6H485
DT 01-WAR-2003 (TREMBLrel. 23, Created)
DT 01-WAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-WAR-2004 (TREMBLrel. 26, Last annotation update)
DE Putative tubby related protein.
GN Name=P0450A04.125;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
RT clone:P0450A04."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004274; BAC20077.1; -; Genomic_DNA.
DR HSSP; P50586; 117E.
DR Gramene; Q6H485; -.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB_1; UNKNOWN 1.
DR PROSITE; PS01201; TUB_2; UNKNOWN 1.
SQ SEQUENCE 406 AA; 45454 MW; 2FD1F9AB24FEB84 CRC64;

Query Match 58.1%; Score 1139; DB 2; Length 406;
Best Local Similarity 56.4%; Pred. No. 4.7e-92;
Matches 232; Conservative 57; Mismatches 86; Indels 36; Gaps 8;

QY 1 MTPRSLQSEM-----RSRPHRVVHAAASTANSSDPPS-----WSELPE 38
Db 1 MSFRSVIQEKGKIGAISRRGRFRGRVRRVRAAAEPPDESSAAALVRRSCWTQLPP 60

QY 39 ELLREILIRVETVDGDFWPSRRNVAVACGYCRSWRLITKEIVAVPFFSSKLTFFISLKOS 98
Db 61 ELLREVLARVESEGW-WPRRDVAVACGYCRSWRGIVREIVRTPEASGNLTFFISLKQP 119

QY 99 GPRDSLVOQCFIKRNRNTQSHLYLGLTSTLDNGKFLAASKLRATCTDYIISLRSDDI 158
Db 120 GPRDAPMKCFIVNRRTTQTYLYLGLTLDALTDGKFLAARKRRTCTEYILSLDMNDI 179

QY 159 SKRSNAYLGRMRNFIQKTFVDSQTG---AKMQKRSRNF1KVSPRVPGSGYP1A 214
Db 180 SKRTDSYVKLSRNF1QKTFVDSQTG---AKMQKRSRNF1KVSPRVPGSGYP1A 239

QY 215 HISYELNVLGSRGPRRMRIMDTIPMSIVSRGVASTSISSPSRSPVPRSHSK 270
Db 240 HISYELNVLGSRGPRRMRIMDTIPMSIVSRGVASTSISSPSRSPVPRSHSK 297
```

```
QY 271 PLRSNSASCSGNNLGPPLVLSNKAPRWHEQLRCWCLNFGHGRVTVASVKNFQLVAVSD 330
Db 298 SYRANNTASLTQN--GSKLVLSNKAPRWHEQLRCWCLNFGHGRVTVASVKNFQLVASDE 355
QY 331 CE-AQGTSERIILQFGKVKDMFTMDYGYPISAFOAFAICLSSFFETRIACE 380
Db 356 SNPTNQEHDDVILQFGKVKDMFTMDYGYPISAFOAFAICLSSFFETRIACE 406

RESULT 7
Q69U54 ORYSA
ID Q69U54_ORYSA PRELIMINARY; PRT; 451 AA.
AC Q69U54
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Putative chain A, C-terminal domain of mouse brain tubby protein.
GN Name=P0015C07.29;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC
RT clone:P0015C07."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004654; BAD33172.1; -; Genomic_DNA.
DR Gramene; Q69U54; -.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR PRINTS; PR01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB_1; UNKNOWN 1.
DR PROSITE; PS01201; TUB_2; 1.
SQ SEQUENCE 451 AA; 50528 MW; F875D1FB652B881E CRC64;

Query Match 51.7%; Score 1015; DB 2; Length 451;
Best Local Similarity 49.8%; Pred. No. 5.2e-81;
Matches 212; Conservative 57; Mismatches 87; Indels 70; Gaps 9;

QY 7 LOEMRSRPHRVVHAAASTANSSDPPFSWSELPEELREILIRVETVDGDFWPSRRNVAVCA 66
Db 44 LSELDRPWW-----DQSRWVGLPELLRDVNMKLEEG-SNWPSSKDVAVCA 91

QY 67 GYCRSWRLITKEIVAVPFFSSKLTFFISLKOSGPRDSLVOQCFIKRNRNTQSHLYLGLT 126
Db 92 AVCRTWREICKDIVQSPFCGKLTFFVSLKQPGPRDGLIQCFIKRDKSKLTLYLCLSP 151

QY 127 S-LTDNGKFLAASKLRATCTDYIISLRSDDIISKRSNAYLGRMRNFIQKTFVETD--- 182
Db 152 AVLSENGKFLAASKLRATCTDYIISLRSDDIISKRSNAYLGRMRNFIQKTFVETDTP 211

QY 183 ----GS----QTGAARQVKRSRNF1KVSPRVPGSGYP1AHISYELNVLGSRGPRMRCI 234
Db 212 PYNAGSLMSCQHSRRISRR-----VSPKLTGSGYPIAHVKYELNVLGTRGPRMQCT 265

QY 235 MDTIPMSIVSRGVAVASTSISSPSRSPVPRSHSKPLR-----SNSASCSDS 282
Db 266 MHSIFASAVDPGCVVPGQPEQLLPGFESFESTNTSSRFSFMDRSLDPSRFSISGS 325

QY 283 GNNLGP-----PLVLSNKAPRWHEQLRCWCLNFGHGRVTVASVKNFQLVAVSDCEAG 334
Db 326 ANQOGEDDIPEAKERPLVLRNKPWHEQLRCWCLNFGHGRVTVASVKNFQLVAVSDCEAG 385

QY 335 QTSE-----RILOFGKVKDMFTMDYGYPISAFOAFAICLSSFE 374
Db 386 QLEQOQOQOQOQHASSSSASDHGKVILOFGKVKDMFTMDYGYPISAFOAFAICLSSFE 445
```


[illegible]


```
AC Q9ST78;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CAA303719.1 protein.
GN Name=q3037.19;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
NCBI_TaxID=4530;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DNA.
RA Hong G., Chen Z.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ245900; CAB53492.1; -; Genomic DNA.
DR HSSP; P50586; I17E.
DR Gramene; Q9ST78; -.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB_1; UNKNOWN_1.
DR PROSITE; PS01201; TUB_2; 1.
SQ SEQUENCE 462 AA; 51513 MW; 456367D114AE59F4 CRC64;

Query Match
Best Local Similarity 50.4%; Score 989; DB 2; Length 462;
Matches 210; Conservative 70; Mismatches 93; Indels 96; Gaps 10;

QY 1 MTFRSLQEMR-----SRPRVVHAAASTANGSDPF-----SNSELP 37
D 1 MSFRSIVRDRDGFGLSGRRGFEVRLGHRKKGSHGAVHELHDPVPVIOSSCWASLPPEL 60
QY 38 EELLREILRVETVGGDWPSPRRNVACAGVCSWRILTKETIVAVPEFSKLTFFISLKQ 97
D 61 PELLRDVWKLEEE-DDSNWPSRKOVVACASVCTTWRDMCKDQIVRNPEFCGLTFFVSLQ 119
QY 98 SGRPSDLVQCFTKRNNTQSYHLGLTTS-LTNGKFLAASKLRATCTDYIISLRSD 156
D 120 PGRDGVICFTKDKSKLTYHLCLSSAVLDTGKFLLSAKRSRTHTHYIISMDSK 179
QY 157 DISKSNAYLGRMSNFLTGTFTVPD-----GSQTGAAMQKSRSSNFTKVSRRVPQGS 211
D 180 NISRSSGGYIGKLSNFLTGTFTFYDTPPNARTLCSQERTSRRPSSRKVPKVPKTCY 239
QY 212 PTAHISYELNVLGSRGPRMRMCIMDTPMSIVESRGVV-----ASTS 253
D 240 PIVQVNYELNVLGTRGPRMRMCAMHSIPASAVEPGGIVFGQPKELLPLRFESFRSMAT 299
QY 254 ISFSRSRSPVFRSHSKPLRSNSASCSDGNNLGL-----DPLVLNSKAPRW 301
D 300 FKYS-----ITDHTDSSSRPSEFGGALQCGEQEQDGDVVKERPLVLNKA 353
QY 302 EQLRCWCLNFHGRVTVASVKNFOLVAVS-----DCEAGQTSER----- 339
D 354 EQLQWCLNFRGRVTVASVKNFOLIAAQPASGAASEPSEQAAQOQQTQPSQSSSS 413
QY 340 -----IILQFGVKGMFTMDYGPISAFQAFALICLSSPETRIACE 380
D 414 SSSSNHDTVILQFGKVAKDMFTMDYRPLSAFOAFALICLTSFDTKLACE 462

RESULT 14
Q53PP5 ORYSA
ID Q53PP5_ORYSA PRELIMINARY; PRT; 440 AA.
AC Q53PP5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Tub family, putative.
GN ORFNames=LOC_Os11g06420;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
NCBI_TaxID=39947;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RA Buell C., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Lin H.,
RA Zhu W., Hamilton J., Jones K., Tallon L., Feldblyum T., Beutin T.,
RA Bera J., Kim M., Jin S., Fadrosh D., Vuong H., Overton II L.,
RA Reardon M., Weaver B., Johri S., Lewis M., Utterback T., Van Aken S.,
RA Wortman J., Haas B., Koo H., Zismann V., Hsiao J., Iobst S.,
RA de Vazeilles A., White O., Salzberg S., Fraser C.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]_
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC120533; AAX95106.1; -; Genomic DNA.
SQ SEQUENCE 440 AA; 48955 MW; 9D9D7312F67AB9B4 CRC64;

Query Match
Best Local Similarity 49.4%; Score 969; DB 2; Length 440;
Matches 207; Conservative 72; Mismatches 93; Indels 76; Gaps 12;

QY 1 MTFRSLQEMR-----RSRPRVVHAAASTANGSDPFWSLSPEEL 40
D 1 MSFRSIVRDRDGFGLSGRRGFEVRLGHRKKGSHGAVHELHDPVPVIOSSCWASLPPEL 60
QY 41 LREILRVETVGGDWPSPRRNVACAGVCSWRILTKETIVAVPEFSKLTFFISLKQSGP 100
D 61 LRDITLEREES-ATWPSRKHVACAGVCRTRWRECKEIVKNPELGGKITFFISLRQGP 119
QY 101 RDSLVQCFTKRNNTQSYHLGLTTS-LTNGKFLAASKLRATCTDYIISLRSDDIS 159
D 120 RDGTWQCFTIRDKSTQTYLYLSLGSVAVDNGKFLLSAKRNWHATCTEYVISMANNLS 179
QY 160 KRSNAYLGRMSNFLTGTFTVPD-----GSQTGAAMQKSRSSNFTKVSRRVPQGS 210
D 180 RSTNTNIGKLSNFLTGTFTFYDTPYNATSDSQSGKT---SRRFSN--KGTAKHPCST 234
QY 211 YPIAHISYELNVLGSRGPRMRMCIMDTPMSIVESRGVVAST-----SIGSFSSRSRP 263
D 235 YSIANISYELNVLGTRGPRMRMCMLHSSIPASSLEAGTVPSQPDLSLAHLSNESSFRSV 294
QY 264 VFRS-----HSKPLRSNSASCSD-----SGNNLG-----DPLVLNSKAPRWHEQL 304
D 295 FSKSSVMDHS--MHFSSAQFSDISIGDGPRIGRVLSDDDECKETPLILQNKA 352
QY 305 RCWCLNFHGRVTVASVKNFOLVAVSDCEAGQTS-----ERILQFGVKGMFTMDY 352
D 353 QCWCLNFRGRVTVASVKNFOLIAAQTAPAGAPTSPQVPPPPPHDKVILQFGKVAKDMF 412
QY 353 TMDYGVPIASAFOAFALICLSSPETRIACE 380
D 413 TMDYHPLSAFOAFALICLSSFDTKLACE 440

RESULT 15
Q8GVES ARATH
ID Q8GVES_ARATH PRELIMINARY; PRT; 394 AA.
AC Q8GVES;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tubby-like protein 2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
NCBI_TaxID=3702;
```

```
[1]
RN NUCLEOTIDE SEQUENCE.
RX PubMed=15064372; DOI=10.1104/pp.103.037820;
RA Lai C.P., Lee C.L., Chen P.H., Wu S.H., Yang C.C., Shaw J.F.;
RT "Molecular analyses of the Arabidopsis TUBBY-like protein gene
  family.";
RL Plant Physiol. 134:1586-1597(2004).
DR EMBL; AY045773; AAK98801.1; -; mRNA.
DR HSSP; PS0586; IIIE.
DR GO; GO:0045449; P:regulation of transcription; TAS.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB_1; UNKNOWN 1.
DR PROSITE; PS01201; TUB_2; UNKNOWN 1.
SQ SEQUENCE 394 AA; 43878 MW; 2BBA39F8666AEDC0 CRC64;

Query Match          49.3%; Score 966.5; DB 2; Length 394;
Best Local Similarity 50.1%; Pred. No. 8.5e-77;
Matches 201; Conservative 64; Mismatches 107; Indels 29; Gaps 7;

QY 1 MTFRSLQEMR-----SRPHVVHAAASTANSDPFS-WSELPPEELREIL 45
DB 1 MSLSILRLDKVRDGLGIGISKESWSKSHIAPDQTTPDNIPOSPWASLPPELLHDII 60

QY 46 IRVETVGGDWPSRRNVACAGVCRSWRLTKETIVAVPFPSSKLTPTPISLKQSGPRDSL 105
DB 61 WRVESEETA-WPARAAVVCASVCKSWRGITMEIVRIPEQCGKLTPTPISLKQSGPRDSPI 119

QY 106 QCFIKENRNTQSYHLYLGLTSLTNGKFLAASKLKRATCTDYIILSRDDISKESNAY 165
DB 120 QCFIKENRATATYILYGLMPSETENDKLLAARRIRATCTDFIILSAKNFSRSSSTY 179

QY 166 LGRMRNFLTGTFTVFDG-SQTAAKQKSRSNFTKVSPPVQSGYPPIAHISYELNLVIG 224
DB 180 VKLRSGLTGTFTIYDNTASTAQAPNRLHPKQAPKLPNTSSTVGNITYELNLVLR 239

QY 225 SRGPRMRIMDTPMSIYVESRGVVASTSISSFSSRSSPVFR--SHSKPLRSNSASCSDS 282
DB 240 TRGPRMHCAKDSIPLSSVIAPSPVVOGIEEVSSSPSPKGETITTTDKETPDNSPS--- 295

QY 283 GNNLGDPLVLGNKAPRWEHQRLRCWCLNFHGRVTVASVKNFQLVAVSDCEAG---QTSER 339
DB 296 ---LRDQPLVLKNKSPRWEHQRLQCWCLNFKGRVTVASVKNFQLVAVSDAPPEEHER 352

QY 340 IILQFGKVGKMDTMDYGYPIAFAQAFACLSSFETRIACE 380
DB 353 VILQFGKIGKIDTMDYRPLSAFAQAFACISSFDTKPACE 393
```

Search completed: December 23, 2005, 23:27:49
Job time : 235 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 23:20:27 ; Search time 46 Seconds
(without alignments)
682.973 Million cell updates/sec

Title: US-10-763-042-9
Perfect score: 1962
Sequence: 1 MFRSLQLQMRSPRHRVHA.....SAFQAPATCLSSFFETRIACE 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCUTS COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	441.5	22.5	506	2	US-09-032-365A-19
2	437.5	22.3	460	1	US-08-630-592-7
3	437.5	22.3	460	1	US-08-714-991-7
4	437.5	22.3	460	2	US-09-032-365A-8
5	437.5	22.3	506	1	US-08-631-200-8
6	437.5	22.3	506	1	US-08-829-553-8
7	437.5	22.3	506	1	US-08-922-267A-8
8	437.5	22.3	506	1	US-08-936-707A-8
9	437.5	22.3	506	1	US-08-936-706A-8
10	437.5	22.3	506	2	US-09-248-203-8
11	437.5	22.3	506	2	US-09-032-365A-62
12	437.5	22.3	506	2	US-08-812-824-4
13	437.5	22.3	506	2	US-09-406-071-8
14	437.5	22.3	506	2	US-08-955-918C-10
15	437.5	22.3	506	2	US-08-697-766A-10
16	437.5	22.3	506	2	US-09-814-986-8
17	437.5	22.3	512	2	US-09-032-365A-60
18	437.5	22.3	518	2	US-09-032-365A-58
19	437.5	22.3	561	1	US-08-714-991-27
20	437.5	22.3	561	2	US-09-032-365A-10
21	436.5	22.2	459	1	US-08-630-592-2
22	436.5	22.2	459	1	US-08-714-991-2
23	436.5	22.2	459	2	US-09-032-365A-2
24	436.5	22.2	505	1	US-08-631-200-2
25	436.5	22.2	505	1	US-08-630-592-4
26	436.5	22.2	505	1	US-08-714-991-4
27	436.5	22.2	505	1	US-08-829-553-2

28	436.5	22.2	505	1	US-08-922-267A-2	Sequence 2, Appli
29	436.5	22.2	505	1	US-08-936-707A-2	Sequence 2, Appli
30	436.5	22.2	505	1	US-08-936-706A-2	Sequence 2, Appli
31	436.5	22.2	505	2	US-09-248-203-2	Sequence 2, Appli
32	436.5	22.2	505	2	US-09-032-365A-4	Sequence 4, Appli
33	436.5	22.2	505	2	US-08-812-824-3	Sequence 3, Appli
34	436.5	22.2	505	2	US-09-406-071-2	Sequence 2, Appli
35	436.5	22.2	505	2	US-08-914-986-2	Sequence 2, Appli
36	427	21.8	504	2	US-08-955-918C-7	Sequence 7, Appli
37	427	21.8	504	2	US-08-697-766A-7	Sequence 7, Appli
38	423	21.6	460	2	US-09-270-767-45069	Sequence 45069, A
39	418	21.3	285	1	US-08-631-200-15	Sequence 15, Appl
40	418	21.3	285	1	US-08-829-553-15	Sequence 15, Appl
41	418	21.3	285	1	US-08-922-267A-15	Sequence 15, Appl
42	418	21.3	285	1	US-08-936-707A-15	Sequence 15, Appl
43	418	21.3	285	2	US-08-936-706A-15	Sequence 15, Appl
44	418	21.3	285	2	US-09-248-203-15	Sequence 15, Appl
45	418	21.3	285	2	US-09-406-071-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-032-365A-19
; Sequence 19, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502th, Michael
; APPLICANT: Nishina, Patsy
; APPLICANT: Naggart, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,365A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-032-365A-19
Query Match 22.5%; Score 441.5; DB 2; Length 506;
Best Local Similarity 36.9%; Pred. No. 1.8e-40;


```

; MOLECULE TYPE: protein
US-08-922-267A-8

Query Match      22.3%; Score 437.5; DB 1; Length 506;
Best Local Similarity 34.9%; Pred. No. 5e-40;
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;

QY 78 EIVAVPEPSSKLTFFPISLQSGPRDSLQVCFIKRNRNTQ-----SYHLVILGLTSLTDN 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 EVQDLEEFALR-----PAQGITIKCRITRDKKGMDRGMYPTFLHLD-----RED 292

QY 132 GK--FLLAASKUKRATCTDYIISLRSDDISKRSNAYLGRMRSNPLGTFKTFVFDGSGQTGAA 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 GKKVFLLAGRKKKSKTSNYLISVDPDTLSRGDSYIGKLSNLMGKTFVYD-----NGV 348

QY 190 KMKSRSSNFIVKSPRVPQGSYPPIAHISYELNVLGSRGPRMRCIMDTP-MSIVESRGV 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 NPQKASSSTLESGLTR-----QELAAVCYETNVLGFGKGRKMSVI---VPGMNMVHER-- 398

QY 249 VASTSISSFSRSSPVFRSHSKPLRSNSASCSDSGNLGDPLVLNKAAPRWHQLRCWC 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 399 ---VSIRPRNEHETLLARQNK-----NTESI-----IELONKTPVWDDTQSYV 440

QY 309 LNFHGRVTVASVKNFQVAVSDCEAGQTSERIILOFGVKGMFTMDYGYPISAFOAPAI 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 441 LNFHGRVTVASVKNFQII-----HGNDPDIYVMQFGRVAEDVFTMDYNYPLCALQAPAI 494

QY 369 CLSSFETRIACE 380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 495 ALSSFDKSLACE 506

RESULT 8
US-08-936-707A-8
; Sequence 8, Application US/08936707A
; Patent No. 5871931
; GENERAL INFORMATION:
; APPLICANT: Kley, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,707A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-936-707A-8

Query Match      22.3%; Score 437.5; DB 1; Length 506;
Best Local Similarity 34.9%; Pred. No. 5e-40;
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;

QY 78 EIVAVPEPSSKLTFFPISLQSGPRDSLQVCFIKRNRNTQ-----SYHLVILGLTSLTDN 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 EVQDLEEFALR-----PAQGITIKCRITRDKKGMDRGMYPTFLHLD-----RED 292

QY 132 GK--FLLAASKUKRATCTDYIISLRSDDISKRSNAYLGRMRSNPLGTFKTFVFDGSGQTGAA 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 293 GKKVFLLAGRKKKSKTSNYLISVDPDTLSRGDSYIGKLSNLMGKTFVYD-----NGV 348

QY 190 KMKSRSSNFIVKSPRVPQGSYPPIAHISYELNVLGSRGPRMRCIMDTP-MSIVESRGV 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 NPQKASSSTLESGLTR-----QELAAVCYETNVLGFGKGRKMSVI---VPGMNMVHER-- 398

QY 249 VASTSISSFSRSSPVFRSHSKPLRSNSASCSDSGNLGDPLVLNKAAPRWHQLRCWC 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 399 ---VSIRPRNEHETLLARQNK-----NTESI-----IELONKTPVWDDTQSYV 440

QY 309 LNFHGRVTVASVKNFQVAVSDCEAGQTSERIILOFGVKGMFTMDYGYPISAFOAPAI 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 441 LNFHGRVTVASVKNFQII-----HGNDPDIYVMQFGRVAEDVFTMDYNYPLCALQAPAI 494

QY 369 CLSSFETRIACE 380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 495 ALSSFDKSLACE 506

RESULT 8
US-08-936-707A-8
; Sequence 8, Application US/08936707A
; Patent No. 5871931
; GENERAL INFORMATION:
; APPLICANT: Kley, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,707A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-936-707A-8

```

```

Best Local Similarity 34.9%; Pred. No. 5e-40;
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;

QY 78 EIVAVPEFSKLTPIISLKQSGPRDSLVCQFIKENRNTQ-----SYHLVILGLTSLTDN 131
Db 247 EVQDLEEFALR-----PAPGQITIKRITRDKKGMGRMGPVTFLLHLD-----RED 292

QY 132 GK--FLAAASKLKRATCTDYIISURSDDISKRSNAYILGRMSNFLGTFKTVFDGSGQTAA 189
Db 293 GKKVFLAGARRKRSKTSNLIYVDPDLSRGDSYIGKLSNLMGKFTVYD-----NGV 348

QY 190 KMKRSRSNFTKVSPRVQGSYPTAHISYELNVLGSRGPRRMCIMDTIP--MSIVESRGV 248
Db 349 NPQKASSTLESGLR-----QELAAVCYEYNVLGPKGPKMSVI---VPGMNWVHER-- 398

QY 249 VASTSISSPSRSPVPSRHSKPLRSNSASCSDSGNNLGDPLVLNKNAPRWHEQLRCWC 308
Db 399 ---VSIRPNEHETLLARWQNK-----NTESI-----IELQNKTPWNDDTQSYV 440

QY 309 LNFHGRVTVASVKNFQLVAVSDCERAGQTSERIILQPKVGKMDMTDYGYPISAFQAPAI 368
Db 441 LNFHGRVTVQASVKNFQII-----HGNDPDYIVMFGVRAEDVFTMDYNYPLCALQAFAI 494

QY 369 CLSSFETRIACE 380
Db 495 ALSFSDSKLACE 506

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Query Match          22.3%; Score 437.5; DB 2; Length 506;
Best Local Similarity 34.9%; Pred.No. 5e-40;
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;

QY      78 EIVAVPEFSSKLTFTPISLKQSGPRDLSVQCPIKRNRTQ-----SYHLVGLGTTSLTDN 131
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      247 EVQDEEFAIR-----PAQCITIKRIYTRDKGMDRGMYPTTFHLHD-----RED 292
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      132 GK--FLLAASKLKRATCTDYIIISLRSDDISKRSNAYLGRMRSNPLGKTFVFDGSGQTGAA 189
      | | | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db      293 GKKVFLLAGRKKKSKTSNYLISVDPDLSRGGBDSYIGKLRSLNMGTKFTYVD---NGV 348
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      190 KMKQRSRSNFIKVSPPVQGSYPIAHISYELNVLGSRGPRMRCIMDTIP-MSIVESRGV 248
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      349 NPQKASSSTLESGLTR-----QELAAVCYETNVLGFKGPRKMSVI---VPGMNMVHER-- 398
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      249 VASTSISSFSSRSPPVFRSHKPLRNSASCSDSGNNLGDPPPLVLSNKAPWHQOLRCWC 308
      | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      399 ---VSIIPRNEHETLLARWQK-----NTESI-----IELQNTFPWNDDTQSYV 440
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      309 LNFHGRVTVASVKNFQLVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPISAFOAPAI 368
      | | | | | | | | | | | : | : | : | : | : | : | : | : | : | : |
Db      441 LNFHGRVTVASVKNFQII-----HGNDPDIYVQFGRVAEDVFTMDYNYPLCALQAFAI 494
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      369 CLSSFETRIACE 380
      | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      495 ALSFSFDSKLACE 506
      | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 11
US-09-032-365A-62
; Sequence 62, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502th, Michael
; APPLICANT: Nishina, Patsy
; APPLICANT: Naggart, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; TITLE OF INVENTION: NEUROSENSORY DEFECTS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,365A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```



```
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 506 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-09-406-071-8
```

Query Match 22.3%; Score 437.5; DB 2; Length 506;
Best Local Similarity 34.9%; Pred. No. 5e-40;
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;

QY 78 EIVAVPEFSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQ-----SYHLVGLTSTLTDN 131
Db 247 EVQDLEEFALR-----PAPOGITTICRIIRDKGMGRMGWPTTYFLHLD-----RED 292

QY 132 GK--FLLAASKLRATCTDYIILSRSDDISKRSNAYLGRMSNFLGTFKFTVFDSQTCAA 189
Db 293 GKGVFLLAGRKRSKTSNYLSVDPTDLRGSDSYIGKLRSNLMTGMTFTFYD---NGV 348

QY 190 KMQKSSRNFIKVSPRPQGSYPPIAHISVELNVLGSRGPRMRCIMDTIP-MSIVESRGV 248
Db 349 NPKASSSLESGLTLR-----QELAAVCYETNVLFKGPRKMSVI---VPGMMVHER-- 398

QY 249 VASTSISSFSSRSSPVFRSHSKPLRSNSASCSDSGNLGDPPPLVLNSKAPRWHEQLRCWC 308
Db 399 ---VSIRPNRHETLLARWQNK----NTESI-----IELQNKTPVMNDTQSIV 440

QY 309 LNFHGRTVASVKNFQVAVSCBAGQTSERIILOFGVKGMFTMDYGYPISAFQAFAI 368
Db 441 LNFHGRTVASVKNFQII-----HGNDPDYIVMOGRVAEDVFTMDYNPLCALQAFAI 494

QY 369 CLSSFETRACE 380
Db 495 ALSSFDKSLACE 506

RESULT 14
US-09-955-918C-10
Sequence 10, Application US/08955918C
Patent No. 6268130
GENERAL INFORMATION:
APPLICANT: Klevn, Patrick, and Moore, Karen
TITLE OF INVENTION: RP Compositions and Therapeutic and
TITLE OF INVENTION: Diagnostic Uses Therefor
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,918C
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/697,766
FILING DATE: 29-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MNI-007CPDV2CPA
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid

```
; LENGTH: 506 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-955-918C-10
```

Query Match 22.3%; Score 437.5; DB 2; Length 506;
Best Local Similarity 34.9%; Pred. No. 5e-40;
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;

QY 78 EIVAVPEFSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQ-----SYHLVGLTSTLTDN 131
Db 247 EVQDLEEFALR-----PAPOGITIICRIITRDKKGMGRMGWPTYFLHLD-----RED 292

QY 132 GK--FLLAASKLRATCTDYIILSRSDDISKRSNAYLGRMSNFLGTFKFTVFDSQTCAA 189
Db 293 GKGVFLLAGRKRSKTSNYLSVDPTDLRGSDSYIGKLRSNLMTGMTFTFYD---NGV 348

QY 190 KMQKSSRNFIKVSPRPQGSYPPIAHISVELNVLGSRGPRMRCIMDTIP-MSIVESRGV 248
Db 349 NPKASSSLESGLTLR-----QELAAVCYETNVLFKGPRKMSVI---VPGMMVHER-- 398

QY 249 VASTSISSFSSRSSPVFRSHSKPLRSNSASCSDSGNLGDPPPLVLNSKAPRWHEQLRCWC 308
Db 399 ---VSIRPNRHETLLARWQNK----NTESI-----IELQNKTPVMNDTQSIV 440

QY 309 LNFHGRTVASVKNFQVAVSCBAGQTSERIILOFGVKGMFTMDYGYPISAFQAFAI 368
Db 441 LNFHGRTVASVKNFQII-----HGNDPDYIVMOGRVAEDVFTMDYNPLCALQAFAI 494

QY 369 CLSSFETRICE 380
Db 495 ALSSFDKSLACE 506

RESULT 15
US-08-697-766A-10
Sequence 10, Application US/08697766A
Patent No. 6399760
GENERAL INFORMATION:
APPLICANT: Gimeno, Carlos; Kleyn, Patrick; and Moore, Karen J.
TITLE OF INVENTION: RP Compositions and Therapeutic and
TITLE OF INVENTION: Diagnostic Uses Therefor
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/697,766A
FILING DATE: 29-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Silveri, Jean M.
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: MNI-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-697-766A-10

Query Match      22.3%; Score 437.5; DB 2; Length 506;
Best Local Similarity 34.9%; Pred. No. 5e-40;
Matches 109; Conservative 60; Mismatches 82; Indels 61; Gaps 12;

QY 78 EIVAVPEESSKLTFFPISLKSGPRDSLVQCFFIKENRNTQ-----SYHLVGLGTTSLTDN 131
Db 247 EVODLEEFALR-----PAQGITIKRITRDKKGMDRGMYPYFLHLD-----RED 292

QY 132 GK--FLLAASKLKRATCTDYIISLRSDDISKRNAVILGRMRSNFLGTFKTFVFDGSGTGAA 189
Db 293 GKKVFLLAGKRRKSKTSNVLISVDPDLSRGGDSYIGKLSNLMGTGKFTVYD-----NGV 348

QY 190 KMQKSRSSNFIKVSVPVQGSYPPIAHISYELNVLGSRGPRMRCIMDTIP-MSIVESRGV 248
Db 349 NPQKASSSTLESGLR-----QELAAVCYETNVLGPKGPKMSVI---VPGMNMVHER-- 398

QY 249 VASTSISSEFSSRSPVFRSHSKPLRSNSASCSDSGNNLGDPPPLVLSNKAAPRWHEQLRCWC 308
Db 399 ---VSIRPRNEHETLLARWQNK-----NTESI-----IELQNKTFVWNDTQSYV 440

QY 309 LNFHGRVTIVASVKNFOLVAVSDCEAGOTSRIILQFGKVKDMFTWYGYPISAFOAFAI 368
Db 441 LNFHGRVTQASVKNFQII-----HGNDPDYIYVQFGRVAEDVFTWYNYPLCALQAFAI 494

QY 369 CLSSPETRIACE 380
Db 495 ALSSPDSKLACE 506
```

Search completed: December 23, 2005, 23:29:23
Job time : 48 secs

THIS PAGE BLANK

Query Match	100.0%;	Score 1962;	DB 5;	Length 380;
Best Local Similarity	100.0%;	Pred. No. 5.6e-192;		
Matches 380;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MTFRSLLOEQRSPHRVYVHAASSTANSSDPFSKSELPEELLREILIRVETVDGGDWPSRR	60	
Db	1	MTFRSLLOEQRSPHRVYVHAASSTANSSDPFSKSELPEELLREILIRVETVDGGDWPSRR	60	
Qy	61	NVVACAGVCSWRLLTKEIVAVPEFFSKLFPISLKQSGPRDSLVQCFIKENRNTQSYHL	120	
Db	61	NVVACAGVCSWRLLTKEIVAVPEFFSKLFPISLKQSGPRDSLVQCFIKENRNTQSYHL	120	
Qy	121	YLGITTSLTNGKFLLAASKLKATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV	180	
Db	121	YLGITTSLTNGKFLLAASKLKATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV	180	
Qy	181	FDGSQTGAAMQKRSRNSNFIKVSPRVQGSYPYIAHISVELNVLGSRGPRRMCIMTDITPM	240	
Db	181	FDGSQTGAAMQKRSRNSNFIKVSPRVQGSYPYIAHISVELNVLGSRGPRRMCIMTDITPM	240	
Qy	241	SIVESRGVASTSISPSRSHSPVFRSHSKPLRNSASCSDSGNNIGDPPLVLNKAAPRW	300	
Db	241	SIVESRGVASTSISPSRSHSPVFRSHSKPLRNSASCSDSGNNIGDPPLVLNKAAPRW	300	
Qy	301	HEQLRCWCLNPHGRVTVASVKNFOLVAVSDCEAGQTSERIIOFGKVGKDMFTWDYGYPI	360	
Db	301	HEQLRCWCLNPHGRVTVASVKNFOLVAVSDCEAGQTSERIIOFGKVGKDMFTWDYGYPI	360	
Qy	361	SAFOAFAICLSSFFETRIACE	380	

```
Db      361  SAFAFAICLSSPETHIACE 380
|||||
RESULT 2
US-10-763-042-11
; Sequence 11, Application US/10763042
; Publication No. US20050014266A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Jai-Fu
; APPLICANT: Lai, Chia-Ping
; TITLE OF INVENTION: PLANT TUBBY-LIKE PROTEINS
; FILE REFERENCE: 08919-099001
; CURRENT APPLICATION NUMBER: US/10/763,042
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 60/441,380
; PRIOR FILING DATE: 2004-01-21
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-10-763-042-11
Query Match      72.1%; Score 1414.5; DB 5; Length 380;
Best Local Similarity 72.9%; Pred. No. 7.9e-136;
Matches 282; Conservative 31; Mismatches 51; Indels 23; Gaps 5;

QY      10  MRSRPHRVH-----AAASTANSDPFSSELPSELLREILIRVETVGGDWPSSRRNVV 63
Db      1  MRSRPHRVHDLAAAAAUSTSVSSQDYRWSELPSELLREILIRVEAAGGGWPSRSVV 60

QY      64  ACAGVCRSWRLTKIIVAVPEFSKLTFFPISLKQSGRDSLVCQCFIKRNRNTQSYHLYLG 123
Db      61  ACAGVCRWRLMNETVWVPEISSKLTFFPISLKQSGRDSLVCQCFIKRNRITQSYHLYLG 120

QY      124  LITSLTDNGKFLAASKLRATCTDYIIISLRSDDIKSRNAYLGRMSNFLTGTFTVFDG 183
Db      121  LITSLTDDGKFLAACKLHTTCTDYIIISLRSDDMSRRSQYVGVKRSNFLTGTFTVFDG 180

QY      184  S---QTGAAMQKRSRNFITKVPYQGSYPIAHISYELNVLSGRGPRMRRCIMDTIPM 240
Db      181  NLLPSTGAALKRSRNPYAKVSLGSPYVAHIITYELNVLSGRGPRMKQCLMDTIP 240

QY      241  SIVESRGVASTSISFSSRSSVPFRSHKPLRSNSASCSDSGNNLGDPLVLNKAAPRW 300
Db      241  STMEPOGVASEPSEFLLGTRSLRSQSKPLRSS-----SHLKETPLVLSNKTPRW 293

QY      301  HEQLRCWCLNFHGRVTVASVKNFOLVAV-SDCEAG-----QTSERIILQFGKVGKDMFT 353
Db      294  HEQLRCWCLNFHGRVTVASVKNFOLVAAGSCSGTGMSPEROSERIILQFGKVGKDMFT 353

QY      354  MDYGYPISAFOAFAICLSSPETHIACE 380
Db      354  MDYGYPISAFOAFAICLSSPETHIACE 380
|||||
RESULT 3
US-10-763-042-3
; Sequence 3, Application US/10763042
; Publication No. US20050014266A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Jai-Fu
; APPLICANT: Lai, Chia-Ping
; TITLE OF INVENTION: PLANT TUBBY-LIKE PROTEINS
; FILE REFERENCE: 08919-099001
; CURRENT APPLICATION NUMBER: US/10/763,042
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 60/441,380
; PRIOR FILING DATE: 2004-01-21
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; ;
```

```
; SEQ ID NO 3
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-10-763-042-3
Query Match      58.6%; Score 1149; DB 5; Length 406;
Best Local Similarity 56.7%; Pred. No. 1.5e-108;
Matches 233; Conservative 58; Mismatches 84; Indels 36; Gaps 9;

QY      1  MTFRSLLOEM-----RSPHRVHAAASTANSDPFSSELPSELL 41
Db      1  MSFKSLIQDMRGELGISIRKGFDFVFGYGRSRSQRVVQDTSVPVDAFKQSCWASMPPELL 60

QY      42  REILIRVETVGGDWPSSRRNVVACAGVCRSWRLTKIIVAVPEFSKLTFFPISLKQSGPR 101
Db      61  RDVLMRIQSE--DTWPSRKNVVSCAGVCRNMRKEIVKBIVRVPELSKLTFFPISLKQSGPR 119

QY      102  DSLVQCFIKRNRNTQSYHLYLGLTSLT-DNGKFLAASKLRATCTDYIIISLRSDDISK 160
Db      120  GSLVQCYIMNRSNQTYIYLGLNQAANDGKFLAACKRFRPCTDYIIISLNCDDVSR 179

QY      161  RSNAYLGRMSNFLTGTFTVFDGSGTQ-AAKQKSRSSNFI---XVSPRVPGSGYPIAH 216
Db      180  GSNYVIGKLRSNFLTGTFTVYDAQPTNPQTQVTRTRSSRLLSLKQVSPRIPSGNYPIVAH 239

QY      217  SYELNVLSGRGPRMRRCIMDTIPMSIVESRGV-----VASTSISFSSRSSVPFRSHSK 270
Db      240  SYELNVLSGRGPRMRRCIMDTIPMSIVESRGV-----VASTSISFSSRSSVPFRSHSK 295

QY      271  PLRSNSASCSDSGNNLGDPLVLNKAAPRWHLQRCWCLNFHGRVTVASVKNFOLVAV-S 329
Db      296  STRAESLSPGSSAAQKEGLLVLNKAAPRWHLQRCWCLNFHGRVTVASVKNFOLVAAPE 355

QY      330  DCEAGQTSERIILQFGKVGKDMFTDYGYPISAFOAFAICLSSPETHIACE 380
Db      356  NGPAGPEHENVILQFGKVGKDMFTMDYQYPISAFOAFAICLSSPETHIACE 406
|||||
RESULT 4
US-10-424-599-149493
; Sequence 149493, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149493
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(391)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106014C.1.pep
US-10-424-599-149493
Query Match      57.7%; Score 1132; DB 4; Length 391;
Best Local Similarity 58.3%; Pred. No. 7.8e-107;
Matches 231; Conservative 63; Mismatches 74; Indels 28; Gaps 11;

QY      6  LIQEMRSRPHRVH-----AAASTANSDPFSSELPSELLREILIRVETVGGDWPSS 58
Db      3  LIRSNRSRQRVVQEQEAAECVGVVGDGSMRQSWANMPQELLREVLRLRIEASE-DTWPP 61
|||||
```


Qy	59	RNVVACAGVCSWRILTKBIIVAVPEFSSKLTFFPISLKQSGPRDULVQCTIKKNRNTQSY	118
Db	62	RKWSVACAGVCSWRHITKDIIVKTPELSSKLTFFISVKQGPRENLLRCPIKKNRSTQTY	121
Qy	119	HYLGLHTTSLTONGKFLLAASKLKRAATCTDVIISLRSDDISKRNAYLGRMRNFIJGTFK	178
Db	122	YLFSLTSLTSLAEDGKFLLAARKCRPTCTDVIISLDADMSKGSVYVKLRNFIJGTFK	181
Qy	179	TVFDGS--OTGAANKQKGRSS---NFIKVSRRVPOG-SYPIAHISYELNV-LGSRGPRRM	231
Db	182	TIYDSQLPTHG-AKTMKSRSTRLVNLKQVSPKVTGXNYPVAHISYEFKXKCLGSRGPRRM	240
Qy	232	RCIMDTPIMSIVBSRGVASTSISFSS-----RSPVFRSHSKPLRNSASCDSGN--N	285
Db	241	HCVMDSIPATAIEPGGVAPTQTEFLSNIDMFPPSPFFRS-----KSNRAENSMGSLVD	295
Qy	286	LGDPPPLVNLKAPRWHEQLRCWCLNFHGRVTVAWSKQFQVLA-VSDCEAGQTSERIILQF	344
Db	296	QKDGMLVLKNAKPRWHEQLQWCLNFHGRVTIASVKNFQVLAWSAENGCPAGPEHDKIILQF	355
Qy	345	GKVGKDMFTMDYGYPISAQPAFACLSSEFTRIACE	380
Db	356	GKVGKOLFTMDYRYPISAFQPAFACLSSEDFTRIACE	391

RESULT 5

US-10-424-599-284656
; Sequence 284656, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 284656
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_9906C.1.pep
US-10-424-599-284656

Query Match	56.2%	Score 1103;	DB 4;	Length 400;
Best Local Similarity	56.3%;	Pred. No. 7.6e-104;		
Matches 232;	Conservative 54;	Mismatches 102;	Indels 44;	Gaps 11;

Qy	1	MTFRSLQEE-----MRSPPHRVVHAASTA-----NSSDPFWSSELPPEELAREIL	45
Db	1	MSPSRITLERKGFVKFGYSMRSRSQSHSDAIAQDSLVLVDGLKQSCWANMPPELLRDVL	60
Qy	46	IRVETVDGGDWPSPRRNVACAGVCRNRILTKETIVAVPEPSSKLTFFPISLKQSPROSLV	105
Db	61	MRLEASE-DSWPAQKQHVACAGVCRNRWKEIKVSPQLSGKUTFPISLKQSPROSLL	119
Qy	106	QCPTKRNRTQS VHYLYGLTSLTLDNGKFLAASKLRATCTDYIISLRSDDIKSRNAY	165
Db	120	QCYIKNRSNQYYLFLGLNQASTDEGKFLLSARKCRATHDTYIISLNCDDVSRGSTY	179
Qy	166	LGRMRSNPLGTKFTVPDG-SQTGAANKQKGRSSNFI---KVSPPVPGSGYPIAHISYELN	221
Db	180	IGKLRSNFLGTKFTVYDAHPPIYCAKVTKSRSTRLSVKQVSPRPAGNYPIAHVSYDLN	239
Qy	222	VLASGRPRMRCIMDTTPMSLIVESRGVVASTSISFSR-----SSPVFRSHS----KPLR	273
Db	240	VLASRGPRIMOCVNDALPASAVEPGG-VAPQTQTFHSRIDTSPSIFFPFRKSTRMNDLP	298

```

Qy 274 SNSASCSDSGNNILGDPPLVLSNKAPRWHEQLRCWCLNFHGRVTVASVKXNPOLVA-----V 328
Db 299 TVP:TCQNEGTF-----LVLRNKSPRWHEHLQCWCLNFHGRVTVASVKXNPOLVASPKNGV 352
Qy 329 SDCEAGOTSERIILOFGKVGKDMFTMDYGYPISAFOAFAICLSSSPETRIACE 380
Db 353 SE-----QAQENVILQFGKVGKDVFTMDYQYPISAFOAFAICLSSSPDTKIACE 400

RESULT 6
US-10-437-963-193011
; Sequence 193011, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with Plant Growth and Development
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 193011
; LENGTH: 407
; TYPE: RPT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(407)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8918C.1.pep
; US-10-437-963-193011

```

Query Match	52.8%	Score	1035.5	DB	4	Length	407
Best Local Similarity	54.6%	Pred. No.	6.5e-97				
Matches	221	Conservative	52	Mismatches	89	Indels	43
Gaps	11						
Qy	1	MTFRSLIQEM-----RSRPH-----RVHAAASTANSSDPFSWSELPEELL	41				
Db	1	MSFRSLIQEMDERGSIISRSHRGGGAPVAAVGPAAAMQOSQWAQLPPELL	60				
Qy	42	REILIRVETVDGGWPSRRNVVACAGVCRSWRLTKGIVAVPEFSKLPFPIISLKQSGPR	101				
Db	61	REVLVRIESEVW-WPSRRDVVACAGVCRSWRGITKEIVRVPEASGLTFPIISLKQGPGR	119				
Qy	102	DSLVCQCFIKRNNTQSHLYLGLTSLTNGKFLAASKLKEATCTDYIISLRSDDISKR	161				
Db	120	DGTLKCFIRRRNTQTYYLYLGLTALADGKFLAARKCRKPTCYDYLISLDMDSWKG	179				
Qy	162	SNAYLGRMRSNFLTGTFTVDGSGQT-GAAQMKSRSSNFI---KVSPPVQGSYPIAHIS	217				
Db	180	SNTYIGKLRSNFLTGTFTYDAHPDYDGAUVVSKSRSAVVGLNQVSPRPAGNYPVSHIS	239				
Qy	218	YELNVLGSRGPRRRCIMDTTPMSIVESRGVVAS-----TSISFSSRSSSPVFRSHSKP	271				
Db	240	YELNVLGARGPRRNCINDSIPTSAVQBGKAPTQTEPFLSGLDSPFSSIS--FFRSKSPR	297				
Qy	272	LRSNSASCSDGNLGDPLVLSNKAAPWHEQRCWCLNFHCG--RVTVASVKNFQLVAVS	329				
Db	298	IEIVQLQLSTQK---EKLVPKNKSPRWHEQLQAWCLNFRGRPRATLASVKNFQL--VS	352				
Qy	330	DCEAGQTSER----IILQFGKVGKDMFTMDYGYPIISAFQAFAILC	370				
Db	353	SDENGPTNOEOKVIXLOFGKIEKDLFTMDYRYPISAFOSPAICL	397				


```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3591-079-H7_FLI_pep
US-10-425-114-72469

Query Match      51.9%; Score 1018; DB 4; Length 436;
Best Local Similarity 48.3%; Pred. No. 4.5e-95;
Matches 213; Conservative 65; Mismatches 97; Indels 66; Gaps 10;

QY 1 MFRSLLOEM-----RSRPHRVVHAAASTANSSDPFSSELPEEL 40
Db 1 MSFRSIVRDVDFGSLSRGFEVKFLGHRGRSGHGAHELDPAPVIOSSCWANLPEEL 60

QY 41 LRILLIRVETDGGWPSRRNVACAGVCRSWILTKETIVAVPEFSKLTFFPISLKOSGP 100
Db 61 LRDIVERLEASE-ATWPNRKSVVSCAAVCKTWREICKETIVKNEPFGSKITFFPVSLLKQGP 119

QY 101 RDSLVQCFIKRNRNTQSYHLYGLTTS-LTDNGKFLLAASKLRATCTDYIISLRDDIS 159
Db 120 RDVTIOCFIKRDKSTQTYIYLYLCLSTAVLVESKFLCAKRTSRPTCTEYTFMNSENLS 179

QY 160 KRSNAYLGRMSNFLTGTFTVFGSOTGAAMQKSRSSNFI--KVSRRVPQGSYPYIAHIS 217
Db 180 RSKMYIGKLSNLLGTKEAIYDTPPCNAABPKTSRRFYSRKVSPKVSSTYNYAQS 239

QY 218 YELNVLGSRGPRMRCIMTIPMSIVESRGVVAS-----TSISFSS-SRSSPVF 265
Db 240 YELNVLGTGPRMNCVMHSIPTSCLEAGTVPQSOLDILARSTDSFGSISFSKSVVD 299

QY 266 RSHSKPLRSNSASCS-----SGNNLQDP-----PLVLSNKAAPRWHEQLRCWCLN 310
Db 300 RS-----MRFSSRFSDISSSRRIGDTASGDNDECKERPLILRNKAPRWHEQLQCWCLN 355

QY 311 FHGRVTVASVKNFQVAVSDCEAGQTS-----BRILQFGKVGKDMFTMDYGYTP 359
Db 356 FGRGVTVASVKNFQVAVSDCEAGQTS-----BRILQFGKVGKDMFTMDYGYTP 415

QY 360 ISAFQAFALCLSSPFTRIACE 380
Db 416 LSAFQAFALCLSSFDTKLACE 436

RESULT 10
US-10-732-923-19498
; Sequence 19498, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 19498
; TYPE: PRT
; LENGTH: 448
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(448)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-19498

Query Match      51.8%; Score 1016; DB 5; Length 448;
Best Local Similarity 49.9%; Pred. No. 7.5e-95;
Matches 206; Conservative 63; Mismatches 100; Indels 44; Gaps 8;

QY 11 RSRPHRVVHAAASTANSSDPFSSELPEELLIRLIRVETDGGWPSRRNVACAGVCR 70
Db 37 RGSQSTVHELCDADLIIOESRWASLPPELLRDLVIRRLSEASTXWPSRKDVVSCAAVCK 96

QY 71 SWRILTKETIVAVPEFSKLTFFPISLKOSGRDLSVOCFIKRNPTQSYHLYGLTTS-LT 129
Db 97 AWRECKEIVLSPFCGKLTFFPLSLAQPGPRDGMIOCFIKRDKSKSYHLYLCLSTAVLA 156

QY 130 DNGKFLLAASKLRATCTDYIISLRDDISKRSNAYLGRMSNFLTGTFTVFD-----G 183
Db 157 DSGKFLLSAKRHKTKTCTEYVISMADNISRSSSTYIGKLSNFLTGTFTFIYDTPSYNG 216

QY 184 SQTGAAMQKSRSSNFIKVSRRVPQGSYPYIAHISYELNVLGSRGPRMRCIMTIPMSIV 243
Db 217 AVIPPPVXRSRRSFRNSKVSXPKMPGSYNYAQVYELNVLGTRGPRMRCVHESIPASSV 276

QY 244 ESRGVV-----ASTSISFSSRSPPVRS---HSKPLRSNSASCSDSGNNL 286
Db 277 EFGGIVPGQPEQIVPAPFESFRSTTSF-SKSIDRSNMGFOXPDPDFSSARSFADIAGGT 335

QY 287 GDP-----PLVLSNKAAPRW-HEQLRCWCLNFGHGRVTVASVKNFQVAVSDCE----- 332
Db 336 SDBEGQNKERPLVLRNKAAPRWHEQLQCWCLNFRGRVTIASVKNFQVAVSDCE----- 395

QY 333 -----AGQTSERILLOFGKVGKDMFTMDYGYPIASFAQAFALCLSSPFTRIACE 380
Db 396 TPSQPAPPQDKIILQFGKVAKDMFTMDYRPLSAFQAFALCLSSFDTKLACE 448

RESULT 11
US-10-732-923-19480
; Sequence 19480, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 19480
; TYPE: PRT
; LENGTH: 455
; ORGANISM: Arabidopsis thaliana
US-10-732-923-19480

Query Match      51.7%; Score 1013.5; DB 5; Length 455;
Best Local Similarity 47.1%; Pred. No. 1.4e-94;
Matches 215; Conservative 64; Mismatches 100; Indels 77; Gaps 10;

QY 1 MFRSLLQEMR-----SRPHRVVHAAASTANSSDPFS-----WSEL 36
Db 1 MSFRSIVRDVDSIGLSRRSDFDKLLSLNKGKSGSVQDSHESQLVVTIQTETPWANL 60

QY 37 PEELLREILIRVETDGGWPSRRNVACAGVCRSWILTKETIVAVPEFSKLTFFPISLK 96
Db 61 PEELLRDVVKLEESE-SWPARRHVACASVCRSDCKEIVQSPESLGGITFFVSLK 119

QY 97 QSGPRDLSVQCFIKRNRNTQSYHLYGLTTS-LTDNGKFLLAASKLRATCTDYIISLRS 155
Db 120 QGPRDATWQCFIKRDKSNLTVHLYLCLSPALLVENGKFLLSAKIRRTTYEYVISMHA 179

QY 156 DDISKRSNAYLGRMSNFLTGTFTVFD-----GSQTGAAMQKSRSSNFI--KVSRRVPQ 208
Db 180 DTISRSSNTYIGKIRSNTFLGTFTFIYDTPAYNSNIARAVQPVGLSRRFYSKRVSPKVS 239

QY 209 GSVPIAHISYELNVLGSRGPRMRCIMTIPMSIVESRGV-----ASTS 253
Db 240 GSYKTAQVSYELNVLGTRGPRMRCAMNSIPASSLAEGTVPQDPIIYVRSILDSFPRS 299

QY 254 ISSFSR-----SSPFRSHSKPLRSNSASCSDSGNNLQDPPLVLSNKAAPRWHEQ 303
Db 300 ITSSSRKITYDYSNDFSSARESDILGPLSEQGVVLESGKRSNPLVLKNKPPRWHEQ 359

QY 304 LRCWCLNFGHGRVTVASVKNFQVAVSDCE-----AGQTS---ERILQF 344
```


US-10-425-115-202431

```
Query Match          51.5%; Score 1010.5; DB 4; Length 437;
Best Local Similarity 48.0%; Pred. No. 2.6e-94;
Matches 212; Conservative 66; Mismatches 97; Indels 67; Gaps 10;

QY 1 MFRSLQRM-----RSRHRVHAAASTANSSDPFWSLPEEL 40
DB 1 MFRSLVRDVRGFGSLRRGFEVFLGHRGRSHGVHLELDPAPVIOSSCWANLPPEL 60

QY 41 LREILIRVTVGGDWPSRRNVVACAGVCRWRILTKIIVAVPEFSSKLTFFPISLKQSGP 100
DB 61 LRDVIERLEASB-ATWPNKSVVSCAAVCRTWREICKEIVKNPEFSGKITFPVSLKQGP 119

QY 101 RSLVQCFIKRNRNTQSYHYLGLTTS-LTDNGKFLLAASKUKRATCTDYIISLRSDDIS 159
DB 120 RDVTIQCFIKRDKSTQTYLYLCLSTAVLVESGKFLLCARSTRPTCTEYTFPMNSNL 179

QY 160 KESNAVLRGRMNFGLTKFTVFDGSGTGAQKQKSRSSNFI--KVSPRVPOQSYPTAHIS 217
DB 180 RSKMVIKGRNLGLTKFAIYDTQPPCNAASPGKTSRRRFYRKVSPKVSSTYINIAQVS 239

QY 218 YELNLVSGRPRMRMRCIMDTIPMSIVESRGVVAS-----TSISFS-SRSSPVF 265
DB 240 HELNLVGTGRPRMNCVMSHSIPTSCLEAGSVPSQLDSLARSTDSEFGSISFSKSSVD 299

QY 266 RSHSKPLRSNSCSD-----SGNNLGD-----PLVLSNKAPRWHEQLRCWCIN 310
DB 300 RS---MRPSSRFSDISMSRRIGDTASGDNDECKERPLILRNKAPRWHEQLQCWCIN 355

QY 311 FHGRVTVASVKVNFOLVAVSDCEAGQTS-----ERILQFGVKDKMTMDYGY 358
DB 356 FHGRVTVASVKVNFOLVAAQPAAGAPTQSOATVPPPEHEKVILOFGKVAKDMFTMDYRY 415

QY 359 PISAFQAFALCLSSPETRIACE 380
DB 416 PLISAFQAFALCLSSFDTKLACE 437
```

RESULT 15

```
US-10-437-963-150007
; Sequence 150007, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 150007
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50284C.1.pgp
US-10-437-963-150007
```

```
Query Match          51.5%; Score 1009.5; DB 4; Length 445;
Best Local Similarity 51.1%; Pred. No. 3.4e-94;
Matches 210; Conservative 66; Mismatches 92; Indels 43; Gaps 11;

QY 11 RSRHRVHAAASTANSSDPFWSLPEELIRIVETVDGDPWPSRRNVVACAGYCR 70
DB 37 RGKQSTVHELCDADLIIOESRWASLPPELLRDVIRRLSE-STWPSRKDVVSCAAVCK 95
```

```
QY 71 SWRIITKEIVAVPEFSSKLTFFPISLKQSGPRDSLVCQCFIKRNRNTQSYHYLGLTTS-lT 129
DB 96 AWRECKEIVLSPFCGKLTFFPLSLKQPGPRDGMICQCFIKRDKSKSTYHYLCLSTAVLA 155

QY 130 DNGKFLLAASKUKRATCTDYIISLRSDDISKESNAVLRGRMNFGLTKFTVFD--GSQTG 187
DB 156 DSGKFLLSAKHRKTKTEYVISMADNISRSSSYIGKLRNFLTGTFFIYDTQPSYNG 215

QY 188 AAKMQKSRSS---NFIKVSPRVPOQSYPTAHISYELNLVSGRPRMRMRCIMDTIPMSIVE 244
DB 216 AVIPPVGRSSRRFNSKVSPPKSPGSYINIAQTYELNLVLTGRPRMHCVMHSIPASSVE 275

QY 245 SRGVV-----ASTSISFSRSPVFRS--HSKPLRSNSCSD-----S 282
DB 276 PGGIIVPGQPEQIVPRAFEESFRSTTSF-SKSSIMDRSMDFSRRDFSSARFSDIAGGTIN 334

QY 283 GNNLG---DPPLVLSNKAPRWHEQLRCWCINLPHGRVTVASVKVNFOLVAV-SDCEAGQTS- 337
DB 335 GDEEGONKERPLVLRNKAPRWHEQLQCWCINLPHGRVTVIASVKNFOLIAAPAPQAGAPT 394

QY 338 -----ERILQFGVKDKMTMDYGYPIISAFQAFALCLSSPETRIACE 380
DB 395 SQPAPPEQDKIILQFGKVAKDMFTMDYRYPLSTFQAFALCLSSFDTKLACE 445
```

Search completed: December 23, 2005, 23:32:33
Job time : 165 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 23:12:06 ; Search time 39 Seconds
(without alignments)
937.497 Million cell updates/sec

Title: US-10-763-042-9
Perfect score: 1962
Sequence: 1 MFRSLQEMRSPHRVHA.....SAFOAFLSSFFETRIACE 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1147.5	58.5	407	2 H84920	probable Tub famil
2	1013.5	51.7	455	2 H96797	hypothetical prote
3	999.5	50.9	445	2 E86382	hypothetical prote
4	929	47.3	388	2 E96513	unknown protein, 3
5	919.5	46.9	386	2 E84562	probable Tub famil
6	787.5	40.1	415	2 F96499	hypothetical prote
7	437.5	22.3	271	2 S42728	phosphodiesterase
8	436.5	22.2	505	2 S68518	tub protein, brain
9	394.5	20.1	425	2 T20691	hypothetical prote
10	352.5	18.0	265	2 T02138	hypothetical prote
11	201	10.2	397	2 E86295	hypothetical prote
12	109.5	5.6	360	2 H86347	hypothetical prote
13	106.5	5.4	695	2 D84634	hypothetical prote
14	103	5.2	322	2 H85068	N7-like protein [i
15	98.5	5.0	391	2 H85063	hypothetical prote
16	98.5	5.0	775	2 T48957	hypothetical prote
17	98.5	5.0	1360	2 T12064	DNA binding protei
18	96	4.9	1161	2 S57180	probable membrane
19	94.5	4.8	1812	2 T49350	breast/ovarian can
20	94	4.8	1532	2 H96795	hypothetical prote
21	93.5	4.8	383	2 P96671	hypothetical prote
22	93.5	4.8	1164	2 S71792	phosphatidylinosit
23	92.5	4.7	736	2 S45859	hypothetical prote
24	92	4.7	542	2 S64030	probable membrane
25	92	4.7	812	2 T52569	squamosa-promoter
26	92	4.7	978	2 H86319	hypothetical prote
27	92	4.7	1035	2 E86342	hypothetical prote
28	91.5	4.7	1295	2 E96549	hypothetical prote
29	90.5	4.6	567	2 JC5957	transforming growt

transforming growt
conserved hypochet
hypothetical prote
ashi protein - fru
flocculation prote
hypothetical prote
hypothetical serin
hypothetical prote
inositol 1,4,5-tri
protein T16B5.8 [i
hypothetical prote
probable transport
acyl-CoA dehydroge
probable carnitine
transforming growt
exoribonuclease 10

30 90.5 4.6 506 2 JC5956
31 90 4.6 560 2 T40608
32 89.5 4.6 1160 2 T23713
33 89.5 4.6 2144 2 S71490
34 89 4.5 1075 2 S48992
35 88.5 4.5 1206 2 E86445
36 88 4.5 1131 2 T41144
37 87.5 4.5 1261 2 T50065
38 87.5 4.5 2833 2 A43360
39 87 4.4 418 2 D86241
40 87 4.4 1286 2 T33476
41 87 4.4 2609 2 T40399
42 86.5 4.4 393 2 AC2892
43 86.5 4.4 393 2 F97667
44 86.5 4.4 579 2 JC5955
45 86.5 4.4 779 2 S64680

ALIGNMENTS

RESULT 1
H84920
probable Tub family protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C:Accession: H84920
R;Lin, X.; Kaul, S.; Rounale, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84920
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-407 <STO>
A:Cross-references: UNIPROT:082257; UNIPARC:UPI00000A8946; GB:AE002093; NID:g3738302; PII
C:Genetics:
A:Gene: At2g47900
A:Map position: 2
C:Superfamily: F-box protein tubby-like, plant type

Query Match 58.5%; Score 1147.5; DB 2; Length 407;
Best Local Similarity 56.6%; Pred. No. 5.6e-92;
Matches 233; Conservative 57; Mismatches 85; Indels 37; Gaps 9;

QY 1 MFRSLQEM-----RSRPHRVVHAHAASTANSSDPFWSGELPEELL 41
Db 1 MSFKSLIQMRGELGISIRKGFDFRGYSRQVRVQDTSPVDFAFKSCWASMPPELL 60
QY 42 REILIRVETVDCGDWPSRRNVACAGVCRSWRLTKEIVAVPVEFFSKLTFPISLKQSGPR 101
Db 61 RDVLWRIEQSE-DTWPSRKNNVSCAGVCNWEIVKEIVRVPVLSKLFPPISLKQSGPR 119
QY 102 DSLVQCFIKRNRNTQSYHLYLGL--TTSLTUNGKFLAASKLRATCTDYIILSRDDIS 159
Db 120 GSLVQCYIMRNRNQTYLYLGLNQAASNDGKFLAARFRFRPTCTDYIILNCDVS 179
QY 160 KESNAYLGMRNFGTKFTVEDGSGTG-AAKQKSRSSNFI---KVSRRVPGSGYPIAH 215
Db 180 RGSNTYIGKLRNFGTKFTVYDAQPTNPGTVTRSRRLSLKQVSPRIPSGNTVPAH 239
QY 216 ISYELNVLGSRPRMRRCIMDTIPMSIVESRGV-----VASTSISFSRSPSPVFRSHS 269
Db 240 ISYELNVLGSRPRMRRCIMDTIPMSIVESRGV-----VASTSISFSRSPSPVFRSHS 295
QY 270 KPLRNSASCSGSDGNLGDPLVLNKNAPRWHEQLRCWCNLFNHRVTVVASVKNFQIVAV- 328
Db 296 KGIARSLPSGSSAAQEGGLVLNKNAPRWHEQLRCWCNLFNHRVTVVASVKNFQIVAAAP 355
QY 329 SCEAQTQSERILLQPGKVGKDMFTMDYGPISAFQAFALCLSSFFETRIACE 380

Db 356 ENGPAGPEHENVLOFGKGVKDVFTMDYQYPISAFOAFTICLSFDTKIACE 407

RESULT 2

H96797

hypothetical protein F22K20.1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004

C:Accession: H96797

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, R.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H96797

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-455 <STO>

A:Cross-references: UNIPROT:Q9ZP59; UNIPARC:UPI000000BEA8; GB:AE005173; NID:g2829918; PIR:PI

C:Genetics:

A:Gene: F22K20.1

A:Map position: 1

C:Superfamily: F-box protein tubby-like, plant type

Query Match 51.7%; Score 1013.5; DB 2; Length 455;

Best Local Similarity 47.1%; Pred. No. 3e-80;

Matches 215; Conservative 64; Mismatches 100; Indels 77; Gaps 10;

QY 1 MTPRSLLQEMR-----SRPHRVVHAAASTANSDPFS-----WSBL 36

Db 1 MSFRIVRDVSIGLSRRSPDFKLSLNKEGKSRGVSQDSHERQLVVTIQTWANL 60

QY 37 PELLREILIRVETVGGDWPSRRNVVACGVCRSWRILTKIIVAVPEFSSKLTTPISLK 96

Db 61 PPELLRDVTKLEESR-SVMPARRHVACASVCRSWRDMCKEIVQSPGLSKITTPVSLK 119

QY 97 QSGPRSLVQCFTKRNNTQSHYLYGLTTS-LTDNGKFLAASKLKRACTDYIISLS 155

Db 120 QPGPRDATQCFTKRDKSNLTHYLYLCLSPALLVENGKFLLSAKRIRRTTYEYVLSMA 179

QY 156 DDISKRSNAYLGRMNFGLTKFTVPD-----GSQTGAAMQKRSRNF--KVSPPVQ 208

Db 180 DTIRSSNTYIGKIRSNELGTPIIYDTPAYNSNIRAVQPVGLSRRFVSKVSPKVS 239

QY 209 GSYPIAHISVELNLVSGRPRMRCIMDTIPMSIVESRGV-----ASTS 253

Db 240 GSYKIAQVSVELNLVGTGRPRMHCAMNSIPASSLAEGGTVPQPDIIVPRSIILDESFRS 299

QY 254 ISFSFSR-----SSPVFRSHKPLRSNSASCSDSGNNLGDPLVLNKNKAPRWHEQ 303

Db 300 ITSSSRKITYDSNDFSSARFSDIIGLUSEDQEVVLEEGKERNSPPLVLKNKPPRWHEQ 359

QY 304 LRCWCLNFHGRVTVASVKNFOLVAVSDCB-----AGQTS--ERILQF 344

Db 360 LQWCLNFRGRVTVASVKNFOLIAANQPQPQPPQPLTQPPSQGTDGDKILLQF 419

QY 345 GKVGKDMFTMDYGYPISAFOAFAICLSFETRIACE 380

Db 420 GKVGKDMFTMDYPRYLSAFOAFAICLSFDTKLACE 455

RESULT 3

E86382

hypothetical protein F4F7.33 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004

C:Accession: E86382

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, R.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86382

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-445 <STO>

A:Cross-references: UNIPROT:Q9FRH7; UNIPARC:UPI000000A1436; GB:AE005172; NID:gl11067277; PIR:PI

C:Genetics:

A:Map position: 1

C:Superfamily: F-box protein tubby-like, plant type

Query Match 50.9%; Score 999.5; DB 2; Length 445;

Best Local Similarity 47.2%; Pred. No. 4.9e-79;

Matches 212; Conservative 71; Mismatches 93; Indels 73; Gaps 11;

QY 1 MTPRSLLQEMR-----RSRPHRV--VHAAASTANSDPFS-----WS 34

Db 1 MSFRIVQDLRDGFGSLRSRSPDFLSSLHKKAQGSFREYSSRDLLSPVIVQTSRWA 60

QY 35 ELPEELLREILIRVETVGGDWPSRRNVVACGVCRSWRILTKIIVAVPEFSSKLTTPIS 94

Db 61 NLPPPELLFDVTKLEESR-SNWPARKHVACASVCRSWRMCQEIIVLGPICGLTTPVS 119

QY 95 LKQSPRSLVQCFTKRNNTQSHYLYGLTTS-LTDNGKFLAASKLKRACTDYIISL 153

Db 120 LKQGPEDAMTQCFTKRDKSKLTFFHLCLSPALLVENGKFLLSAKRTRTRTYEYLSM 179

QY 154 RDDTSKRSNAYLGRMNFGLTKFTVPD-----GSQTGAAMQKRSRNF--IKVSPPV 206

Db 180 DADNLSRSNSVYGLKLRNLFGLTKLVYDTPPPNTSSALITDTSRFRHSRRVSPKV 239

QY 207 PGGSYPIAHISVELNLVSGRPRMRCIMDTIPMSIVESRGV-----249

Db 240 PGGSYNAQITVELNLVGTGRPRMHCIMNSIPISLSLEPGSGVPNQPEKLVAPYSLDSD 299

QY 250 -----ASTSISFSRSPVFRSHKPLRSNSASCSDSGNNLGDPLVLNKNKAPRWHEQLR 305

Db 300 FRNSITFSKSPDHRSLDFSSRFSEM---GISCDNNEEASFRPLILKNKQPRWHEQLQ 356

QY 306 CWCLNFHGRVTVASVKNFOLVAV-----SDCEAGQTS-----ERILQFGKVGKDM 351

Db 357 CWCLNFRGRVTVASVKNFOLVAVARQPOGTGAAAPTSAHPHQDKVILQFGKVGKDM 416

QY 352 FTMDYGYPISAFOAFAICLSFETRIACE 380

Db 417 FTMDYRYPLSAFOAFAICLSFDTKLACE 445

RESULT 4

E96513

unknown protein, 3155-1759 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004

C:Accession: E96513

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, R.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96513
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <STO>
A:Cross-references: UNIPROT:Q9C6B4; UNIPARC:UPI00000A18D9; GB:AE005173; NID:g10092408; P
C:Genetics:
A:Map position: 1
C:Superfamily: F-box protein tubby-like, plant type

Query Match 47.3%; Score 929; DB 2; Length 388;
Best Local Similarity 50.8%; Pred. No. 5.5e-73;
Matches 197; Conservative 65; Mismatches 92; Indels 34; Gaps 12;

QY 11 RSRPH-----RVHAAASTANSDDPSWSELPELLEILIRIVETVGGDWPSRRNVVAC 65
DB 17 RGRSHIAPEGSSVSSLSLTNEGLNQSIWDLPELLDIIQRIES-EQSLWPGRRDWWAC 75
QY 66 AGVCRSWRLITKEIVAVPESSKLTTPISLKQSGPRDSLVOCEIKENRNTQSYHLVGLT 125
DB 76 ASVCKSWREMTKEVVKVPELSGLITTPISLROQPRDAPICQCKRERATGIYRLVGLUS 135
QY 126 TSIT-DNGKFLAASKRATCTDYIISLRSDDISKRSNAYLGRMSNPLGTFTVFDGS 184
DB 136 PALSGDKSKLLSAKVRATGAEFVVSLSGNDFSRSSNYIGKLSNPLGTFTVYENQ 195
QY 185 QTG-AAAMQKRSRNFIVKSPRV--PQGSYPPIAHISYELNVLSGRGPRMRMCIMDTIPMS 241
DB 196 PPPFNKRLPPS-----MQVSPWYSSSSSYNTASILYELNVLRTRGPRMRMCIMHSIPIS 250
QY 242 IVESRGVASTISISSPSSSSPVFRSHSKPLRSNSASCSDS--GNLGDPPPLVLSNKA 299
DB 251 AIOEGGKIQSP--TEFTNQ-----KKKKPLMD---FCSGNLGGESVKEIPELILKNKSPR 301
QY 300 WHEQLRCWCLNFHGRVTVASVKNFQVAVSDCEAG-----QTSERIILQFGKVGKDMF 352
DB 302 WHEQLQCWCLNFGRVTVASVKNFQVAVAA--AAGKGNMIPBEEQDRVILQFGKIGKDI 360
QY 353 TWDYGYPISAFAQFAICLSSFFETRIACE 380
DB 361 TWDYRYPISAFQFAICLSSFFETKPACE 388

RESULT 5
E84562
Probable Tub family protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C:Accession: E84562
R:Bin, X.; Kaul, S.; Rounleley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84562
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-386 <STO>
A:Cross-references: UNIPROT:Q9ZPW1; UNIPARC:UPI00000A83F0; GB:AE002093; NID:g4309738; P
C:Genetics:
A:Map position: 2
C:Superfamily: F-box protein tubby-like, plant type

Query Match 46.9%; Score 919.5; DB 2; Length 386;
Best Local Similarity 49.4%; Pred. No. 3.7e-72;
Matches 198; Conservative 59; Mismatches 107; Indels 37; Gaps 8;

QY 1 MTFRSILQEMR-----SRPHRVHAAASTANSDDPS--WSELPELLEIL 45
DB 1 MSLKSLRLDLKEVRDGLGIGSKRSWSKSHIAPDQTPPLDNIQSPWASLPPELLHDI 60

A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96513
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <STO>
A:Cross-references: UNIPROT:Q9C6B4; UNIPARC:UPI00000A18D9; GB:AE005173; NID:g10092408; P
C:Genetics:
A:Map position: 1
C:Superfamily: F-box protein tubby-like, plant type

Query Match 47.3%; Score 929; DB 2; Length 388;
Best Local Similarity 50.8%; Pred. No. 5.5e-73;
Matches 197; Conservative 65; Mismatches 92; Indels 34; Gaps 12;

QY 11 RSRPH-----RVHAAASTANSDDPSWSELPELLEILIRIVETVGGDWPSRRNVVAC 65
DB 17 RGRSHIAPEGSSVSSLSLTNEGLNQSIWDLPELLDIIQRIES-EQSLWPGRRDWWAC 75
QY 66 AGVCRSWRLITKEIVAVPESSKLTTPISLKQSGPRDSLVOCEIKENRNTQSYHLVGLT 125
DB 76 ASVCKSWREMTKEVVKVPELSGLITTPISLROQPRDAPICQCKRERATGIYRLVGLUS 135
QY 126 TSIT-DNGKFLAASKRATCTDYIISLRSDDISKRSNAYLGRMSNPLGTFTVFDGS 184
DB 136 PALSGDKSKLLSAKVRATGAEFVVSLSGNDFSRSSNYIGKLSNPLGTFTVYENQ 195
QY 185 QTG-AAAMQKRSRNFIVKSPRV--PQGSYPPIAHISYELNVLSGRGPRMRMCIMDTIPMS 241
DB 196 PPPFNKRLPPS-----MQVSPWYSSSSSYNTASILYELNVLRTRGPRMRMCIMHSIPIS 250
QY 242 IVESRGVASTISISSPSSSSPVFRSHSKPLRSNSASCSDS--GNLGDPPPLVLSNKA 299
DB 251 AIOEGGKIQSP--TEFTNQ-----KKKKPLMD---FCSGNLGGESVKEIPELILKNKSPR 301
QY 300 WHEQLRCWCLNFHGRVTVASVKNFQVAVSDCEAG-----QTSERIILQFGKVGKDMF 352
DB 302 WHEQLQCWCLNFGRVTVASVKNFQVAVAA--AAGKGNMIPBEEQDRVILQFGKIGKDI 360
QY 353 TWDYGYPISAFAQFAICLSSFFETRIACE 380
DB 361 TWDYRYPISAFQFAICLSSFFETKPACE 388

RESULT 5
E84562
Probable Tub family protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C:Accession: E84562
R:Bin, X.; Kaul, S.; Rounleley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84562
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-386 <STO>
A:Cross-references: UNIPROT:Q9ZPW1; UNIPARC:UPI00000A83F0; GB:AE002093; NID:g4309738; P
C:Genetics:
A:Map position: 2
C:Superfamily: F-box protein tubby-like, plant type

Query Match 46.9%; Score 919.5; DB 2; Length 386;
Best Local Similarity 49.4%; Pred. No. 3.7e-72;
Matches 198; Conservative 59; Mismatches 107; Indels 37; Gaps 8;

QY 1 MTFRSILQEMR-----SRPHRVHAAASTANSDDPS--WSELPELLEIL 45
DB 1 MSLKSLRLDLKEVRDGLGIGSKRSWSKSHIAPDQTPPLDNIQSPWASLPPELLHDI 60

QY 46 IRVETVGGDWPSRRNVVACVCRSWRLITKEIVAVPESSKLTTPISLKQSGPRDSL 105
DB 61 WRVSEETA-WFAAAVSCASVCKSWRGITMEIVRIPEQCGKLTTPISLKQSGPRDSPI 119
QY 106 QCFIKENRNTQSYHLVGLTTSLTONGKFLAASKLKRATCTDYIISLRSDDISKRSNAY 165
DB 120 QCFIKENRATATVILYGLMPSETENDKLLAARRIRATCTDFFISL-----SSLL 171
QY 166 LGRMSNPLGTFTVFDG--SQTGAAMQKRSRNFIVKSPRVQSGYPPIAHISYELNVLG 224
DB 172 ILMDRSGLGTFTTYIDNQTASTAQAPNRLHPKQAAPKLPNTSGNITVYELNVLR 231
QY 225 SGRPRMRMCIMDTIPMSIVESRGVASTISISSPSSSSPVFR--SHSKPLRSNSASCSDS 282
DB 232 TRGPRMRMCIMDTIPMSIVESRGVASTISISSPSSSSPVFR--SHSKPLRSNSASCSDS 287
QY 283 GNNLGDPPPLVLSNKAAPRHEQLRCWCLNFHGRVTVASVKNFQVAVSDCEAG---QTSR 339
DB 288 ---LRDQPLVKNKSPRHEQLQCWCLNFGRVTVASVKNFQVAVSDCEAG---QTSR 344
QY 340 IILQFGKVGKDMFTWDYGYPISAFAQFAICLSSFFETRIACE 380
DB 345 VILQFGKIGKIDFTWDYRYPISAFQFAICLSSFFETKPACE 385

RESULT 6
F96499
Hypothetical protein T10P12.9 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
C:Accession: F96499
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, N.P.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96499
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-415 <STO>
A:Cross-references: UNIPROT:Q9XIF9; UNIPARC:UPI00000A5392; GB:AE005173; NID:g5080765; P
C:Genetics:
A:Map position: 1
C:Superfamily: F-box protein tubby-like, plant type

Query Match 40.1%; Score 787.5; DB 2; Length 415;
Best Local Similarity 43.2%; Pred. No. 1.3e-60;
Matches 168; Conservative 65; Mismatches 87; Indels 69; Gaps 9;

QY 33 WSELPELLEILIRIVETVGGDWPSRRNVVACVCRSW---RIILTKEIVAVPESSKL 89
DB 55 WANLPAALLRDVWKXLDSE--STWPA-----RSMFYSLLPPDIMSV-----I 96
QY 90 TTPISLKQSGPRDSLVOCEIKENRNTQSYHLVGLTTS--LTONGKFLAASKLKRATCTD 148
DB 97 LTTLLMVKPGRDGIITQYIKRDKSNMTYHLVLSLSPAILVSGKFLLSAKGRSRRATY 156
QY 149 YIISLRSDDISKRSNAYLGRMSNPLGTFTVFD-----GSQTGAAMQKRSRNFIVK 203
DB 157 YVIMADNISRSSSYIGKLSNPLGTFTFYDTAPANNSSQILPGRNRSFNKSKVS 216
QY 204 PRVPGQSYPIAHISYELNVLSGRGPRMRMCIMDTIPMSIVESRGVASTISISSPSS 263
DB 217 PKVPSGSYNIAQVYELNVLLGTGPRMRMCIMHSIPSLALEPGCTVP-----SQP 266

Db 231 AKVRSNALGCTQFTVYDSGQ-----NPKKTTNHAIRQE-----LAAVIYETNVLGFK 277
QY 227 GPRMRCIMDTTPMSIVESRGVAVSTSISSFSRSSPVPFRSHSKPLRSNASCSDSGNL 286
Db 278 GPRKMTVMP-----GIBPPT-----ENRPAVRCPPVRPQDKHTLLERYRLND 320
QY 287 GDPPLVLSNKA PRWHEQLRCWCLNPHGRVTVASVKNFQLVAVSDCEAGOTSERIILQFGK 346
Db 321 LUSLKLSNKS POWNDETQSYLNFHGRVTVQASVKNFQIHQS-----SPEYIVMQFGR 374
QY 347 VGRDMFTMDYGYPISAFOAFAICLSSFPETRIA 378
Db 375 ISDDEFTMDRYPLSAVQAFAIAMSFTHGKLA 406

RESULT 10
T02138
hypothetical protein F8K4.13 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
R:Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Kwan, A.; Yu, G.; Oji, O.; Liu, S.; Li, rtz, D.; Li, J.H.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N submitted to the EMBL Data Library, August 1998
A:Description: Arabidopsis thaliana chromosome 1 BAC F8K4 sequence.
A:Reference number: Z14574
A:Accession: T02138
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-265 <VYS>
A:Cross-references: UNIPROT:O80699; UNIPARC:UPI00000AA02A; EMBL:AC004392; NID:G3282170;
C:Genetics:
A:Gene: ATSP.F8K4.13
A:Map position: 1
A:Introns: 61/3; 110/1; 199/3

Query Match 18.0%; Score 352.5; DB 2; Length 265;
Best Local Similarity 29.3%; Pred. No. 5e-23;
Matches 105; Conservative 38; Mismatches 76; Indels 139; Gaps 12;

QY 36 LPEELRLILRIVETVDGSDWPSRRNVACAGVCSRWRLITKEIVAVPEFSSKLTPTPISL 95
Db 1 MPPELLRDVLMIERSE-DTWPSRKVNVSCVGVCKWNRQIFKEIVNVPVSSKFTPTISL 59
QY 96 KGSGPRDSLVQCFTIKRNNTQSYHLGLTSLTDNGKFLAASKLKRATCTDYIISLRS 155
Db 60 KQPGGSLVQCVKVRNSQTFYLYG----- 87

QY 156 DDISKRSNAYLGRMRNFTLGTETVFDGSGTGAAMKQKRSNFIKVSFPRVPGSGYPIAH 215
Db 88 -----GEAKI-----FCQSEFSEPNKS----- 104

QY 216 ISYELNLVLSGRPRRMRCIMDTTPMSIVESRGVAVSTSISSFSRSSPVPFRSHSKPLRSN 275
Db 105 -----TWKLSLKGPGTRATTOTELDNFVSFRSP----- 131

QY 276 SASCSDSGNLGDPLVLNKA PRWHEQLRCWCLNPHG-RVTVASVKNFQLVAVSDCEAG 334
Db 132 -----SGQKEG--VLVLSKVPRLREQ--SWCLDFNGWRDVS SGGKQLVALL----- 176

QY 335 QTSERILQFGVKGMFTMDYGY-----PISAFQAPA-----IC-----LSSPETRIA 378
Db 177 RTNLRMKTTFSSLRKSETCTNNSYEAWIPLVTSVFAVIARVCRDKGHTPSYELKLA 234

RESULT 11
E86295
hypothetical protein T24D18.17 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
C:Accession: E86295
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86295
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <STO>
A:Cross-references: UNIPROT:Q9S9M8; UNIPARC:UPI00000A0AF61; GB:AE005172; NID:96587813; PII C:Genetics:
A:Map position: 1
C:Superfamily: tubby

Query Match 10.2%; Score 201; DB 2; Length 397;
Best Local Similarity 24.4%; Pred. No. 1.4e-09;
Matches 87; Conservative 54; Mismatches 103; Indels 112; Gaps 18;

QY 70 RSWRILITKE---IVAVPEFSSKLTPTPISLKQSGPRDSL------QCFIKRNRT 115
Db 106 KTWTSVDSEHSSSLKWEFSDSEAPASSWSTLPNRAALLCKTLPDVGRCCTCLIVKEQSP 165

QY 116 QS-----YHLX-----LGLTSLTNGKPLL-AASKLKRATCTDYIISLRSDD 157
Db 166 EGLSHGSVYSLYTHEGRGRKDRKLA VAVHSRRNGKSI FRAVQNVKGLCS-----SDE 218

QY 158 ISKRSNAYLGRMRNFTLGTETVFD-GSOTGA-AKMOKRSNFIKVSFPRV--QGSYPI 213
Db 219 -----SYVGSMTANLLGSKYIYWDKGVGVSGVKNRPLLSSVLI-FTPTITWTGSI-- 269

QY 214 AHISYELNLVLSGRPRRMRCIMDTTPMSIVESRGVAVSTSISSFSRSSPVPFRSHSKPLR 273
Db 270 -----RMRTLL-----PKQPMQKNKKQVQ 291

QY 274 SNSASCSDSGNLGDPLVLVS-----NKAPRWHEQLRCWCLNF--HGRVTV-----ASVKNF 323
Db 292 QASKLPDLWLENKEKIQKLCSPHYNKISKQHE-----LDFRDRGRTGLRIQSSVKNF 345

QY 324 QLVAVSDCEAGOTSERIILQFGVKGMFTMDYGYPISAFOAFAICLSSFPETRIAC 379
Db 346 QULTT-----ETPRQTILQMGVRDKARYVIDFRYPFGYQAFICLASIDSKLCC 395

RESULT 12
B86347
hypothetical protein F24J8.5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: B86347
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86347
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <STO>
A:Cross-references: UNIPROT:Q9LPL4; UNIPARC:UPI00000A85E3; GB:AE005172; NID:99454572; PII C:Genetics:
A:Map position: 1

Query Match 5.6%; Score 109.5; DB 2; Length 360;

Best Local Similarity 26.7%; Pred. No. 0.11; Mismatches 43; Conservative 21; Mismatches 50; Indels 47; Gaps 8;

QY 33 WSELPEELLREILIRVETVDGDPSPRRNVVACAGVCRSWRI-----LTK----- 77
DB 28 WKDIPVELLISLVDD-----RNVIVASGCTGCRDAISFGLTRLRLSCNNNM 78
QY 78 ---EIVAVPEFSSKLTFFPISLKQSGPR--DSLVQCFIKENRNTQSVHLYGLTTSITDNG 132
DB 79 NSLVSLVPKFKVLOT--LNLQDKRPQLEDNAVEAIANHCHELQ--ELDLKSLSKITDRS 134

QY 133 KFLLA-----ASKLRATCTDYIISLRSDDISKRNSAYLGR 168
DB 135 LVLAHGCDDLTKLNLSCGTSP-----SDTAIAYLTR 166

RESULT 13
D84634
hypothetical protein At2g24250 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84634
R:Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84634
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-695 <STO>
A:Cross-references: UNIPROT:Q9ZU40; UNIPARC:UPI000017A78D; GB:AE002093; NID:g4115385; PI
C:Genetics:
A:Gene: At2g24250
A:Map position: 2

Query Match 5.4%; Score 106.5; DB 2; Length 695;
Best Local Similarity 20.4%; Pred. No. 0.5; Mismatches 81; Conservative 52; Mismatches 153; Indels 111; Gaps 17;

QY 4 RSLQEMRSRPHRVHAAASTANSSDPFSWSELPEELLREILIRVETVDGDPSPRRNVV 63
DB 309 QSAVKDLGESSMAMKKISTIMP-DWSQLPEELLHIISTHLE-----DHYFDVAV 360
QY 64 ACAGVCRSWR-----ILTKETAVPEP----- 87
DB 361 HARSVCRSWRSTPPFSSLLRQSYSLPAPPLESKDLCTLEKVLPLFRVLTPPDAADAS 420
QY 88 KLTFFPISLKQSGPRD-----SLVQCFIKEN-RNTQ-----SVHLVGLT 125
DB 421 SEYFLGGLGQDKSNHVELPSPQLQSVKVNVPGETPILNMLDCQIIPIGHKYRLMIGN 480

QY 126 -----TSLTNGKFLAASKLKRATCTDYIISLRSD-----ISKRSNA----- 164
DB 481 PEEYSAAFLPLNEQGGGFEVALLD-----CTDLFLVLRSTEMRWIRLEKSTASCCEL 534
QY 165 --YLGMRGNFLTKFTVDGSGTGAAMQKRSNNFIKVSPPVQGSYPIAHISVELNV 222
DB 535 FTRGRFYATFFNGDTFVIDPSSLEATPLTPHIDSNFL-----VPSGNEELFLV--KTDF 587

QY 223 LSGRGPWRMCIMDTIPMSIVSRGVVASTSISSF--SSRSSVPFRSHKPLRSNASC 280
DB 588 LRCKVSLDBEAAEWVEVSDGLDRVLFLGGLHGNFYCSAKELP----HCGGLTGDILPT 643

QY 281 DSGNNLGDPL--VLSNKA PRWHEOLRCWCLNFHGRV 315
DB 644 VGRNVTYPKYGVHTNKR--KAEDNINCWRSSREN RV 679

RESULT 14
H85068
N7-like protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2004
C:Accession: H85068
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85068
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-322 <STO>
A:Cross-references: UNIPROT:Q9MOU7; UNIPARC:UPI00000A04CC; GB:NC_001268; NID:g7267308; P
C:Genetics:
A:Gene: AT4G05480
A:Map position: 4
C:Superfamily: F-box containing protein

Query Match 5.2%; Score 103; DB 2; Length 322;
Best Local Similarity 20.9%; Pred. No. 0.34; Mismatches 56; Conservative 54; Mismatches 104; Indels 54; Gaps 14;

QY 2 TFRSL-----QEMRSRPHRVHAAASTANSSDPFSWSELPEELLREILIRVETVDGDPW 57
DB 6 TLQSLLMKEDBEORNK-RRTTSTMFLKKDDEERINWDLPELTTISILRLSVTDILD-N 63
QY 58 SRNNVACVAGCRSWRIITKE-----IVAVPEFSSKLTFFPISLKQSG----- 99
DB 64 ARK-----LCRRWRICKDPFMSWRKINLRDCLMAYEFDPSMCHRIVDLSOGGLEINI 116
QY 100 ---PRDSLVOCFIKENRNTQSVHLYGLTTSITDNGKFLAASKLKRATCTDYIISLRSD 156
DB 117 EHFVSDLSLVIVDRSCNLKS--LGISIVPEWTKNG-----VMNGIEKLPLETLVIFHSS 170
QY 157 ---DISKRSNA--YLGMRSNFLTKFTVDGSGTGAAMQKRSNNFIKVSPPVQGSY 211
DB 171 IKDLKAIGHACPOLKTLKLSLGSSEL-AHDISQVGYTIPLEEC-DDDALAIAESMPK--- 225

QY 212 PIAHISYELNVLSRGRPRRMRCTMDTIP 239
DB 226 -LRHLQLMGNGLTWG---LNAILDGCP 249

RESULT 15
H85063
hypothetical protein AT4g05080 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2004
C:Accession: H85063
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85063
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <STO>
A:Cross-references: UNIPROT:Q9S9T0; UNIPARC:UPI00000A6B39; GB:NC_001268; NID:g7367267; P
C:Genetics:
A:Gene: AT4g05080
A:Map position: 4
C:Superfamily: hypothetical protein containing F-box domain

Query Match 5.0%; Score 98.5; DB 2; Length 391;
Best Local Similarity 20.1%; Pred. No. 1.1; Mismatches 62; Conservative 49; Mismatches 116; Indels 81; Gaps 13;

QY 35 ELPEELLREILIRVETVDGDPSPRRNVVACVAGVCRSWRIITKEIVAVPEFSSKLTFFIS 94
DB 6 DLTQDLVKEILSRVITSIG-----AVRSTCKGNALSKDRI----- 42

QY 95 LKQSGPRDSLVOCFIKENRNTQSVHLYLG-----LTTSLTDNGKFLAASKLKR 144
DB 43 LCKAKPKQFHQGFMLSDYRLSRMNFNISTGTFKENGEEFVNLSVKIEGNFLNKRVEISHMY 102

```
QY 145 TCTDYIISLRSDDISKRSNAYLGRMSNFLGTYFTVFGSQTCAAKQKRSRNSNFIKVSP 204
Db |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
103 YCGILLCVTTDTRLVWNPYLGQIRWIQLKTE-TWY---STFCLRYDNNKNHKILRFLD 158
QY 205 RVPQGSYPPIAHISYELNVLGSRGPRMRCTMDTIPMSIVESRGVAVASTSISFSSRSSPV 264
Db ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
159 N-KQGSYEI-----YDLKSYSWRA-----FDVIPKWDIDDDG-----QSASVKGNTY 199
QY 265 FRSHKPLRSNSASCSD-SGNNLG---DPPLVLSNKAPRWHEQLRCWCLNFHGRVTVASV 320
Db ||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
200 FRTIDE--TPNLLICPFTAERFGLDPPFQ-----HGMMSLSWV 238
QY 321 KNFQLVAV 328
Db |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
239 REEKLVAL 246
```

Search completed: December 23, 2005, 23:28:33
Job time : 43 secs

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 03:28:09 ; Search time 1104 Seconds

(without alignments)
8539.032 Million cell updates/sec

Title: US-10-763-042-20

Perfect score: 1140

Sequence: 1 agacgttcgaagttact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.Main.*

- 1: /cgn2_5/prodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_5/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_5/prodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_5/prodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_5/prodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_5/prodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_5/prodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_5/prodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_5/prodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_5/prodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1140	100.0	1140	8	US-10-763-042-20
2	567.8	49.8	1140	8	US-10-763-042-22
3	396	34.7	396	3	US-09-770-423-879
4	367.8	32.3	1629	7	US-10-425-114-8367
5	347.6	30.5	2161	7	US-10-424-599-6651
6	342.6	30.1	1218	8	US-10-763-042-14
7	340.6	29.9	1908	7	US-10-424-599-141814
8	326	28.6	469	3	US-09-770-444-216
9	290.4	25.5	1910	7	US-10-437-963-90528
10	250.4	22.0	1182	8	US-10-763-042-13
11	229.2	20.1	1002	7	US-10-425-114-19311
12	228.6	20.1	1939	7	US-10-424-599-112097
13	220.4	19.3	2358	8	US-10-425-115-119315
14	220.2	19.3	1183	7	US-10-767-701-12426
15	219.8	19.3	2451	7	US-10-424-599-8380
16	210.2	18.4	1164	8	US-10-763-042-17
17	203.6	17.9	2181	7	US-10-424-599-105685
18	195.4	17.1	1335	8	US-10-763-042-21
19	195.4	17.1	1338	6	US-10-225-068-129
20	195.4	17.1	1338	7	US-10-374-780A-2895
21	195.4	17.1	1338	9	US-10-225-068-129
22	193	16.9	1137	8	US-10-763-042-18
23	192	16.8	1878	7	US-10-424-599-64212

24	184	15.1	1287	8	US-10-763-042-16	Sequence 16, Appl
25	174.4	15.3	1357	7	US-10-437-963-57823	Sequence 57823, A
26	169.8	14.9	1910	7	US-10-425-114-22666	Sequence 22666, A
27	169.8	14.9	1930	7	US-10-425-114-22784	Sequence 22784, A
28	169.6	14.9	2253	7	US-10-437-963-47524	Sequence 47524, A
29	165.2	14.5	520	7	US-10-424-599-9312	Sequence 9312, Ap
30	165	14.5	1740	7	US-10-437-963-19517	Sequence 19517, A
31	164.8	14.5	1414	7	US-10-425-114-29388	Sequence 29388, A
32	164.8	14.5	1578	7	US-10-424-599-97364	Sequence 97364, A
33	162	14.2	1926	7	US-10-425-114-26712	Sequence 26712, A
34	162	14.2	2154	8	US-10-425-115-17768	Sequence 17768, A
35	160.8	14.1	1750	7	US-10-425-114-30141	Sequence 30141, A
36	160.8	14.1	1826	7	US-10-437-963-50806	Sequence 50806, A
37	160.4	14.1	571	7	US-10-021-323-4660	Sequence 4660, Ap
38	160.4	14.1	1030	8	US-10-767-795-1188	Sequence 1188, Ap
39	160.4	14.1	2238	7	US-10-437-963-97195	Sequence 97195, A
40	157.8	13.8	707	7	US-10-424-599-6654	Sequence 6654, Ap
41	154.6	13.6	1959	7	US-10-437-963-25577	Sequence 25577, A
42	153	13.4	1269	8	US-10-425-115-35180	Sequence 35180, A
43	148.4	13.0	1365	8	US-10-763-042-12	Sequence 12, Appl
44	147.4	12.9	971	7	US-10-424-599-69809	Sequence 69809, A
45	146.4	12.8	1599	7	US-10-425-114-22378	Sequence 22378, A

ALIGNMENTS

RESULT 1

US-10-763-042-20
; Sequence 20, Application US/10763042
; Publication No. US20050014266A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Jai-Fu
; TITLE OF INVENTION: PLANT TUBBY-LIKE PROTEINS
; FILE REFERENCE: 08919-099001
; CURRENT APPLICATION NUMBER: US/10/763,042
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 60/441,380
; PRIOR FILING DATE: 2004-01-21
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
US-10-763-042-20

Query Match		100.0%	Score 1140;	DB 8;	Length 1140;
Best Local Similarity		100.0%	Pred. No. 0;		
Matches 1140;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGACGCTCCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTCAGCC	60		
DB	1	ATGACGCTCCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTCAGCC	60		
QY	61	GCGGCTCAACCGCTAATAGTTCAGACCTTTTCAGTGTGAGGCTCCGAGGAGCTG	120		
DB	61	GCGGCTCAACCGCTAATAGTTCAGACCTTTTCAGTGTGAGGCTCCGAGGAGCTG	120		
QY	121	CTTAGAGAAATCTGNTAGGCTTGAGACTGTTGACGGCGGCGATTTGGCGTCCGCGCA	180		
DB	121	CTTAGAGAAATCTGNTAGGCTTGAGACTGTTGACGGCGGCGATTTGGCGTCCGCGCA	180		
QY	181	AACGTGGTGGCTTGTCGCGCGCTTTCTCTAGCTGAGGATTTCTACCAAGAGAGATTGTA	240		
DB	181	AACGTGGTGGCTTGTCGCGCGCTTTCTCTAGCTGAGGATTTCTACCAAGAGAGATTGTA	240		
QY	241	GCTGTTCTCGAATTCCTCTAAATTCATTTCCCTATCTCCCTCAAGAGTCTGTTCCA	300		
DB	241	GCTGTTCTCGAATTCCTCTAAATTCATTTCCCTATCTCCCTCAAGAGTCTGTTCCA	300		
QY	301	AGAGATTCCTAGTTCAATGCTTTTAAACGTAATCGAAATCTCAATCGTATCATCTC	360		

Db 988 GGNATGTCACCGAGAGGAGGAGCGAGCGGATATATATGAGTTTGGGAAAGTCGGAAA 1047
QY 1048 GACATGTTACCATGAGATTATGATATCCGATTTCTGCGTTTCAACGGTTTGTCTATCTGC 1107
Db 1048 GATATGTTCCAGTATGATATGATATCCGATTTCTGCGTTTCAACGGTTTGTGCAATTTGC 1107
QY 1108 CTGAGCAGTTTTCAGAACCAAGATTGCTCTGTGAA 1140
Db 1108 TTGAGCAGCTTTGAGAGTAGAATCGCTGTGTGAA 1140

RESULT 3

US-09-770-423-879
; Sequence 879, Application US/09770423
; Publication No. US20020040490A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kriker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2028 (PALA-017PRV)
; CURRENT APPLICATION NUMBER: US/09/770,423
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,512
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 879
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

Query Match 34.7%; Score 396; DB 3; Length 396;
Best Local Similarity 100.0%; Pred. No. 3e-124;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 160 GCGATTTGCGCGTGGCGGAAACGTTGGTGTGTCGCGCGGTTTGTCTAGCTGAGG 219
Db 1 GCGATTTGCGCGTGGCGGAAACGTTGGTGTGTCGCGCGGTTTGTCTAGCTGAGG 60
QY 220 ATTCTCACCAGAGATTTAGCTGTTCTGTAATTCCTCTCTAAATTTGACTTTCCCTATC 279
Db 61 ATTCTCACCAGAGATTTAGCTGTTCTGTAATTCCTCTCTAAATTTGACTTTCCCTATC 120
QY 280 TCCTCAAGCAGTCTGGTCCAGAGATTTCTAGTTCAATGCTTTTATAAAACGTAATCGA 339
Db 121 TCCTCAAGCAGTCTGGTCCAGAGATTTCTAGTTCAATGCTTTTATAAAACGTAATCGA 180
QY 340 AATFACTCAATCGTATCATCTCTATCTCGGATTAACCTACCTCTTTGACGGAACCGGAAG 399
Db 181 AATFACTCAATCGTATCATCTCTATCTCGGATTAACCTACCTCTTTGACGGAACCGGAAG 240
QY 400 TTTCTTCTGCTGCTTCTAAGCTGAAGCGGGCACTTGGCACTGATATCATCTCTTTG 459

Db 241 TTTCTTCTTGTGCTTCTTAAGCTGAAGCGGCAACTTGCACATGATATCATCTCTTTG 300
QY 460 CGTTCCAGCAGATATCTCAAGAGAGCAACGGTATCTTTGGGAGATGAGATCGAACTTC 519
Db 301 CGTTCCAGCAGATATCTCAAGAGAGCAACGGTATCTTTGGGAGATGAGATCGAACTTC 360
QY 520 CTTTGGAAACAAATTCACGGTCTTTTGTATGATGATGATGAG 555
Db 361 CTTTGGAAACAAATTCACGGTCTTTTGTATGATGATGATGAG 396

RESULT 4

US-10-425-114-8367
; Sequence 8367, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8367
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700749863_FLI
US-10-425-114-8367

Query Match 32.3%; Score 367.8; DB 7; Length 1629;
Best Local Similarity 63.2%; Pred. No. 3.2e-114;
Matches 674; Conservative 0; Mismatches 367; Indels 26; Gaps 6;
QY 95 GCTGTGCGGAGCTCCCGAGAGAGCTGTAGAGAAATCTGTAGTTGAGGTTGAGACTGTG 154
Db 156 GTTGGGCCAACATGCCCAAGAGCTTCTCCGAGAGGTCTCTCCGAATCGAG--GCCT 212
QY 155 ACGGCGCGATTTGGCGTCCGCGGAAACGTTGGTGGCTTTGCGCGCGTTTGTCTAGCT 214
Db 213 CCGAGGACACGTGGCGCGCGGAAAGAGCGTCTCTCTCGCGCGCGCTCTGCCGAGCT 272
QY 215 GGAGGATTTCTCACCAGAGATTTAGCTGTTCTCGAATTCCTCTTAATTTGACTTTTC 274
Db 273 GGAGACATATCACCAGAGATTTGCAAAACGCCCGAATCTCTCTCAAGATCACCTTCC 332
QY 275 CTATCTCCCTCAAGCAGTCTGGTCCAGAGATTTCTAGTTCAATGCTTTTATAAAACGTA 334
Db 333 CCAATTTCTTTAAACAGCTGCGCCCAAGGAGAAATCTCTACGGTGTCTTTATAAGCGCA 392
QY 335 ATCGAAATACTCAATCGTATCATCTCTATCTCGGATTAACCTACCTCTTTGACGGAATCG 394
Db 393 ACCGGTCCACCCAAACATCTATTTGTTTCTCAGTTTAAACAGTAGCTAGCTGAAGATG 452
QY 395 GGAAGTTTCTTCTGCTGCTTAAGCTGAAGCGGCAACTTGCACATGATATCATCTCT 454
Db 453 GGAAGTTTCTTCTGCTGCTTAAGCTGAAGCGGCAACTTGCACATGATATCATCTCT 512
QY 455 CTTTGGTTTCACAGCATATCTCAAGAGAGCAACGCGTATCTTTGGGAGAAATGAGATCGA 514
Db 513 CTCCTTGAACGATGATGTCAAGGGAGCAATCTCTATGTTGGGAAACTTAAGATCAA 572
QY 515 ACTTCTTGGAAACAAATTCACGGTCTTTTGTATGATGATGATGATGATGATGATGATG 571
Db 573 ACTTCTTGGAAACAAATTCACAAATCTATGATGATGATGATGATGATGATGATGATG 632
QY 572 TCGAGAGAGCGCGCTCTTC-----TAATTTTCATAAAGTTTTCACCTAGAGTTCTCTC 622

Db 633 TTATGAAAGTGGCTCCACAGCGTGGTGAATCTAAAGCAAGTTTCCACCAAGGTTCCCTA 692
QY 623 AGGGAAGTTACCCCAATCGCTCAATTTCAATGAGTTAAACGTTTACGCTCTCGGGGAC 682
Db 693 CAGGCAATATCCAGTGGCCCAATATTTCAATGAATGAATGCTAGGCTCAAGGGGTC 752
QY 683 CGAGAGAAATCGTTCATCATGATGATACATACCTATGAGCATCGTGGAGTCGAGAGAG 742
Db 753 CTAGGAGAAATGCAATGTGTGATGATGATTCATTCCTGCCCATGCAATGTAACCTGGAGGTG 812
QY 743 TAGTAGCTTCAACATCCATAAGCTCTTTTCCAGTGGCTCATCCAGGCTTTTAGGTCTC 802
Db 813 TAGCACCTACAGAC--TGAGTTTCTCTTAACAACATAGACATGTTTCTTCATTCCTC 870
QY 803 ACTCAAAACCAATGGCCAGTAAATAGTGCATCATGTAGCGACTCAGG-----CAACAACC 856
Db 871 CTTTTTTTTCGATCAAAATCAATCGTGGGAAAATTCATGCTGGACCAATGGTTGATC 930
QY 857 TGGAGATCCACCATTTGGTGTCTGAGCAACAAGCTCCACGGTGGCATGAGCAAGTTACGTT 916
Db 931 AAAAGGATGGGATGCTAGTGTGTAAGAAACAAGGCCCTTAGGTGGCATGAGCAAGTGCAT 990
QY 917 GCTGGTGTCTTAAATTTCCATGTTGCTGAGTCAAGTGGCTTCGGTTAAGAACTTTTACGCTTG 976
Db 991 GTTGGTCCCTGAACCTTTCAATGAGCGGTGACAAATGCCCTAGTTAAAACTTTTACGCTGG 1050
QY 977 TGGCAGTTAGTG---ACTGTGAAGCAGGGCAGACATCTGAGAGGATCATCTCCAGTTTG 1033
Db 1051 TGGCTTCTGCAGAAAACGGACCTGTGGACCAAGAACACGATAGATCATCTCCCAATTTG 1110
QY 1034 GGAACTTGGGAAGGACATGTTTACCATGATGATGATGATGATGATGATGATGATGATGATG 1093
Db 1111 GAAAAGTTGGGAAGGATTTGTTTACAATGATGATGATGATGATGATGATGATGATGATG 1170
QY 1094 CGTTTCTATCTGCTGAGCAGTTTGAACCAAGAAATGCTGTGAA 1140
Db 1171 CATTTGCAATCTGCTCAGCAGTTTGTGATACCAAGATGCTGTGAA 1217

RESULT 5

US-10-424-599-6651
; Sequence 6651, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 6651
; LENGTH: 2161
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106014C.1
US-10-424-599-6651

Query Match 30.5%; Score 347.6; DB 7; Length 2161;
Best Local Similarity 63.2%; Pred. No. 3.2e-107;
Matches 676; Conservative 0; Mismatches 364; Indels 30; Gaps 8;
QY 95 GCTGGTGGAGCTCCCGAGGAGCTGCTTAGAGAAATCTGATAGGTTGAGACTGTTG 154
Db 387 GTTGGGCAACATGCGCCCAAGAGCTTCTCCGAGAGGTCCTCTCCGAAATCGAG--GCCT 443
QY 155 ACGCGCGGATGCGCGTCCGCGGAAACGTTGGCTGGCTGGCGGCTTTTGTGAGCT 214

RESULT 6

US-10-763-042-14
; Sequence 14, Application US/10763042
; Publication No. US20050014266A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Jei-Fu

Db 444 CCGAGGACACGTGGCGCCGCGGGAAGAGCGTCTCTCTCGCGCGCGCTCTCGCGCAGCT 503
QY 215 GGAGATTTCTCACCAGGAGATTTAGCTGTTCTCTGAATTTCTCTCTCTCTCTCTCTCTCTCT 274
Db 504 GGAGACATATCACCAGGAGATTTGTCACAAACGCGCGAACTCTCTCTCTCTCTCTCTCTCT 563
QY 275 CTATCTCCTCTCAAGCAGTCTGTGTCACAGAGATTTCTCTAGTTTCAATGCTTTTATAAACGTA 334
Db 564 CCAATTTCTGTTAAACAGCTGGCCCAAGGGAATCTCTCTACGGTGTCTTTATAAGCGCA 623
QY 335 ATCGAAATPACTCAATCGTATCATCTCTATCTCGGAATTAATCACTCTTTTGAACGATAACG 394
Db 624 ACCGGTCCACCAACATCTATTTGTTCTCAGTTTAAACAGTACGCTAGCTGAGATG 683
QY 395 GGAAGTTTCTTCTGCTGCTTCTAAGCTGAAGCGCGCACTTGGACTGATGATCATCATCT 454
Db 684 GGAAGTTCTTCTGCTGCAACAAATGCAGACGCTCCAACTGACACAGATTTATATCATCT 743
QY 455 CTTTCGGTTCAGACATATCTCAAGAGAGCAACGCGTATCTTGGGGAATGAGATCGA 514
Db 744 CTCCTGACGCAATGATGATGTCACAAAGGAGCACTCTCTATGTTGGGAACTTAAGATCAA 803
QY 515 ACTTCTTTGGAAACAAATTCACGGTCTTTGATGGTGTAGTACAGAC---CGAGAGCGGAAGA 571
Db 804 ACTTCTTTGGAAACAAATTCACAACTATGATAGCAGCTGCTCATACAGAGCAAGA 863
QY 572 TCCAGAGAGCGGCTCTTC-----TAATTTTCAATCAAGTTTCACTACCTAGATTCCT- 621
Db 864 TTATGAAAGTGGCTTCCCAAGGCTGGTGAATCTAAAGCAAGTTTTCACCAAGGTTCTCTA 923
QY 622 CAGGGAAGTTACCCCATCGCTCACATTTTATACAGAGTT--AAACGCTTTAGGCTCTCGGG 679
Db 924 CAGGCTAATCTCAGTGGCCCATTTTCAATGATGATGATGATGATGATGATGATGATGATG 983
QY 680 GACCAGAGAGATGGGTTGTCATCATGATGATGATGATGATGATGATGATGATGATGATG 739
Db 984 GTCTAGGAGATGCAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1043
QY 740 GAGTAGTAGTTCAACATCCATAGCTCTTTTTCAGTGC-----GTCACTCACAGTCT 793
Db 1044 GTGTAGTCTCTACACAGACTGAGTTTCTCTTAAACAACATAGACATGTTTCTCTCATTC 1103
QY 794 TTAGGTCTCACTCAAAACCATTTGCGCAGTAAATGATGATGATGATGATGATGATGATGATG 853
Db 1104 CTTTTTTTTCGATCAAAATCAAAATCTGCGGAAATTCATGCTGAGCACTGTTGATC 1163
QY 854 ACCTGGAGATCCACCATTTGGTGTGTCAGCAACAAAGCTCCACGGTGGCATGAGCACTTAC 913
Db 1164 AAAAG--GATGGGATGCTAGTGTGTTGAAAAACAAGGCCCTTAGTGGCATGAGCAGCTGC 1220
QY 914 GTTGTGTTGCTTAAATTTCCATGTTGTCAGTGCACAGTGGCTTCGGTTAAGAACTTTTACG 973
Db 1221 AATGTTGTTGCTGCTGAACTTTTCAATGAGCGGTTGACAAATTCGCTCAGTTAAAACTTTT 1280
QY 974 TTGTGTCAGTTAGTG---ACTGTGAAGCAGGCGCAGACATCTGAGAGGATCATCTCCAGT 1030
Db 1281 TGGTGGCTTCTGCAAGAAACCGACCTGCTGGACCAAGAACACGATAGATCATCTCCAAT 1340
QY 1031 TTGGGAAAGTTGGGAAGGACATGTTTACCATGATTTATGATATCCGATTTTCTGCGTTTC 1090
Db 1341 TTGGGAAAGTTGGGAAGGATTTGTTTACAATGATTTACCGGTACCTATCTCGGCATTT 1400
QY 1091 AAGCGTTTCTATCTGCTGAGCAGTTTGTGAACCAAGAAATTCGCTGTGA 1140
Db 1401 AGGCATTTTGAATCTGCTCAGCAGTTTGTGATACCAAGATTTGCTGTGAA 1450


```
QY 455 CTTTGCGTTGACGAGATATCTCAAGAGAGCAACGCGTATCTTGGGAGATGAGATCGA 514
Db 839 CTCCTAACTGTGATGATGATCAAGAGGAGTAGTACCTATATGTAAGAGTTGAGATCAA 898
QY 515 ACTTCTCTGGAAACAAAATTCACGGTCTTTTGATG-----GTAGTCAGACCGGAG 562
Db 899 ACTTCTTGGCCACCAATTCACAGTGTATGATGCACACCCCTCAATTTTATGGAGCCAAAG 958
QY 563 CAGCGAAGATGACGAAGAGCGCTCTTCTAAATTTATCAAAAGTTTACCTAGATTCCTC 622
Db 959 TTACAAGTCTCGTTCCACGAGCTAGTTAGTCTCAAGCAAGTTTCTCCAAGAGTTCTCG 1018
QY 623 AGGGAGTTACCCATCGCTCAATTTTCATACGAGTTAAACGTTTAAAGCTCTAGGCTCTCGGGAC 682
Db 1019 CTGGCAACTATCCCATTTGCTCATGTGTCTATGATCTGAATGTTTGGGCTCTAGGGCC 1078
QY 683 CGAAGAAGATGGGTTGTCATCATGATGATACATACCTATGAGCATCGTGGAGTCGCGAGGAG 742
Db 1079 CTAGATAATGCATGTTGTTATGGATGCCATCCCTGCCCTCAGCTGTTGAACCTGGAGGTG 1138
QY 743 TAGTAGCTTCAACATCCATAAGCTCTTTTTCAG-----TCGTCATCACCAGTCTTTTA 796
Db 1139 TGGCCCAACACAGACTCAATTTCTTATAGCAGAAATGATATCTTCTCCATCCATCCCTT 1198
QY 797 GGTCTCACTCAAAACCATTTGGCAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACC 856
Db 1199 TCTTTAGATCAAAATCAACCCGATGGCAATCTCCCA---ACAGTACCTTTGACTTTGTC 1255
QY 857 TGGGAGATCCACCATTTGGTGTGAGCAACAAAGCTCCAGGTGGCATGAGCAGTTACGTT 916
Db 1256 AATATGAGGGACACTGGTATTAGAAACAAGTCCCAAGTGGCATGAACACCTTCAAT 1315
QY 917 GCTGTGCTTAAATTTCCATGTGCGAGTCAAGTGGCTTCGGTTAAGAACTTTACGCTTG 976
Db 1316 GCTGTGTCTGAATTTCAATGGCGAGTGACAGTTGCTTCAGTTAAATAATTTCCAGCTGG 1375
QY 977 TGGCAGTTAGTACTGTGAAG---CAGGGCAGACATCTGAGAGGATCATCTCCAGTTTG 1033
Db 1376 TTGCTTCTCCCAAAAATGGAGTTTCTTGAGCAGGCTCAGGAAAATGTAATTTCTACAGTTTG 1435
QY 1034 GGAAGTTGGGAGGACATGTTTACCATGGATTTAGGATATCCGATTTCTGGCTTTCAAG 1093
Db 1436 GAAAAGTTGGAAGGATGATTTCCACATGGATTTACGATATCCAACTCTGCTCTTCAAG 1495
QY 1094 CGTTTGTCTATCGCTCAGCAGTTTTCGAAACCAAGATTCGCTGTGAA 1140
Db 1496 CATTTGCAATATCGCTTAGCAGCTTTTGACACCAAGATTTGCTTGTGAA 1542
```

RESULT 8

```
US-09-770-444-216/c
; Sequence 216, Application US/09770444
; Patent No. US2002023280A1
```

GENERAL INFORMATION:

```
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
```

```
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE OF INVENTION: thaliana
; FILE REFERENCE: 2027 (PARA-016PRV)
; CURRENT APPLICATION NUMBER: US/09/770,444
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,502
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 216
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-444-216
```

Query Match 28.6%; Score 326; DB 3; Length 469;

Best Local Similarity 100.0%; Pred. No. 3.3e-100;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 815 TCGCAGTAATAGTCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACCATTTGG 874
Db 469 TCGCAGTAATAGTCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACCATTTGG 410
QY 875 TGCTGAGCAACAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGGTCTTAAATTTCC 934
Db 409 TGCTGAGCAACAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGGTCTTAAATTTCC 350
QY 935 ATGCTGAGTCAAGTGGCTTAAAGAACTTTTTCAGCTTGTGGCAGTTAGTACTGTG 994
Db 349 ATGCTGAGTCAAGTGGCTTAAAGAACTTTTTCAGCTTGTGGCAGTTAGTACTGTG 290
QY 995 RAGCAGGCGACATCTCAGAGGATCATCTCCAGTTTGGGAAAGTTGGGAAGGACATGT 1054
Db 289 RAGCAGGCGACATCTCAGAGGATCATCTCCAGTTTGGGAAAGTTGGGAAGGACATGT 230
QY 1055 TTACCATGGATTTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTCCCTGAGCA 1114
Db 229 TTACCATGGATTTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTCCCTGAGCA 170
QY 1115 GTTTTGAACCAAGAAATTCCTGTGAA 1140
Db 169 GTTTTGAACCAAGAAATTCCTGTGAA 144
```

RESULT 9

US-10-437-963-90528

```
; Sequence 90528, Application US/10437963
; Publication No. US20040123343A1
```

GENERAL INFORMATION:

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
```

```
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 90528

LENGTH: 1910

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_8918C.1

US-10-437-963-90528

Query Match

25.5%; Score 290.4; DB 7; Length 1910;

[illegible]

Db 1859 TCAATCATTTGCCAATCTGTCTGAG 1882

RESULT 10
US-10-763-042-13
; Sequence 13, Application US/10763042
; Publication No. US20050014266A1
; GENERAL INFORMATION:
; APPLICANT: Shaw, Jai-Fu
; APPLICANT: Lai, Chia-Ping
; TITLE OF INVENTION: PLANT TUBBY-LIKE PROTEINS
; FILE REFERENCE: 08919-099001
; CURRENT APPLICATION NUMBER: US/10763,042
; CURRENT FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: US 60/441,380
; PRIOR FILING DATE: 2004-01-21
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
US-10-763-042-13

Query Match 22.0%; Score 250.4; DB 8; Length 1182;
Best Local Similarity 55.4%; Pred. No. 4.8e-74;
Matches 585; Conservative 0; Mismatches 441; Indels 30; Gaps 4;

Qy	97	TGGTCCGAGCTCCGCGAGGAGCTGCTTAGAGAAATCCTGATTAGGGTTGAGACTGTTTCAC	156
Db	142	TGGGCTTCTTTGCGCCTGAGTTGCTTCATGACATTAATCTGGAGGGTTGAAGAGATGA-	200
Qy	157	GGCGGGGATTTGGCCGTCGCGCGGAAACGTTGGTGGCTTTGTGCCGCGTTTGTGCTGAGCTGG	216
Db	201	--GACAGCTTTGGCCCGCTCGAGCTGCCGTTGTCTCTTTGTGCTTCAGTATGTAATCATGG	258
Qy	217	AGGATTTCTCACCAGGAGATTGTAGCTGTTCCTGAAATCTCCCTCTAAATTGACATTTCCCT	276
Db	259	AGAGGAATCACTATGGAGATTGTGAGGATCCCTGAGCAGTGTGGGAAGCTCACATTTTCCA	318
Qy	277	ATCTCCCTCAAGCAGTCTGCTCCAAGAGATTCTCTAGTTCAATGCTTTATAAAACGTAAT	336
Db	319	ATCTCATGAAACAGCGCGGCTCGAGACTTCCAATTCANATGTTTTATTAAGAGGAC	378
Qy	337	CGAATAACTCAATCGTATCATCTCTATCTCGGATTAACACTCTTTTGAACGATAACGGG	396
Db	379	AGAGCAACAGCTACATACATTTCTATATGTTGATGCTTCGGAGACTGAGAACGAC	438
Qy	397	AAGTTTCTTCTGTGCTTCTAAGCTGAAGCGGCACTTGACCTGATTAATCATCTCTCT	456
Db	439	AAACTGTTGTTAGCAGCAAGAAGATTAGAAGAGCGCATGTCACAGACTTTTATAATCTCC	498
Qy	457	TTGGGTTTCAGACGATATCTCAAGAGAGCAACCGCTATCTTTGGAGAGTACAGATCGAAC	516
Db	499	CTATCTGCAAGAACTTCTCAGGAGCAGCAGTACTTATGTTGGCAATTAAGGCTCTGGT	558
Qy	517	TTCTCTTGGAAACAAATTCACGGTCTTTTG---ATGGTAGTCAGACCGGAGCGAAGATG	573
Db	559	TTTCTGGGAACCAAGTTCAAAATATAGCAACCAACAGCATCATCCACAGCAACAGCC	618
Qy	574	CAGAAGAGCGCTCTCTTAATTTTCATCAAAAGTTTTCACCTAGAGTTCTCTCAGGAGTTAC	633
Db	619	CAACCTAACCGAAGACTCCACCCGAAACAAAGCGGCTCTTAACTACCTACGAAATAGCTCT	678
Qy	634	CCCATCGCTCACATTTTCATACGAGTTAAACGTCTTAGGCTCTCGGGGACCGAGAGAATG	693
Db	679	ACCGTAGGAACATTAACCTACGAGCTCAATGTTCTTCGCAACAGGGGACCTAGAGAATG	738
Qy	694	CGTTGTCATCGGATACAAATACCTATGAGCATCTGTGGAGTCGCGAGGAGTAGTAGCTTCA	753
Db	739	CAGTGGCTATGGATTCTATACCCCTCTCTCTGTTTATGCTGAAACCGCTCAGTAGTTCAA	798
Qy	754	ACATCCATAAGCTCTTTTCCAGTCGGTCAATCAACAGCTTTTATAGGTCATCAAAAACA	813

[illegible]

RESULT 13

RESOL 13
IIS-10-425-115-119315

US-10-423-113-119313
: Sequence 119315. Application US/104251115

; sequence 11313, Application US
; Publication No. US20040214272A1; FUDIFICATION NO: 0520
: GENERAL INFORMATION:

```

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 119315
; LENGTH: 2358
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_40299C.1
US-10-425-115-119315

```

Query Match	19.3%; Score 220.4; DB 8; Length 2358;
Best Local Similarity	55.2%; Pred. No. 1.4e-63;
Matches	562; Conservative 0; Mismatches 411; Indels 45; Gaps 5;
Qy	168 GCCTGCGCGCAAAAGTGTGCTTGTGCGGGGTTTGTGCTAGCTGGAGGATTCTCAC 227
Db	
Qy	1065 GCGGCGCGCGCGCAGGTGTGGCATGCGCAGCCATTTCGGGTCTGTGGCGGAGGTTCAC 1124
Db	
Qy	228 CAAGGAGATTGTAGCTGTTCTCCTGAAATCTCCTCTAAATTGACTTTCCTATCTCCCTCAA 287
Db	
Qy	1125 CAAGGAGTGGTGAAGACGCTCGAGGAGTGTGGCAGGATTACCTTCCCATATGCCCTTAA 1184
Db	
Qy	288 GCAGTCTGGTCCAAGAGATTCTCTAGTTCGAATGCTTTTATAAAGCTAATCGAATACTCA 347
Db	
Qy	1185 GCAGCGGGGGCCCTCGTGAATGCGGTCCTAGTGTGTTGTGAGGAGAGAGGGCAACATC 1244
Db	
Qy	348 ATCGTATCATCTCTATCTCGGAATTAACACTCTCTTTGA---CGGATAACGGGAAGTTTCT 404
Db	
Qy	1245 CACATATCTACTCTACTTAGGGCTCAGCCCATCTATGAATGTGGAGAAACAGAATTATT 1304
Db	
Qy	405 TCTTGTCTGCTTTAAGCTGAAGCGCGCAACTTGGACHTGATTACATCATCTCTTTTGGCTTC 464
Db	
Qy	1305 GCTTGTCTGCCGGAAGGTCAAGGGGTGCGCAGGAGACTTATTTGTGTATCATCTGGTCTC 1364
Db	
Qy	465 AGACGATATCTCAAGAGAGCAACGGGTATCTTGGGAGATGAGATCGAACTTCCTTGG 524
Db	
Qy	1365 TGAATGATTTCTCTCATCTCTTAGTAGCACTACGTTTGGCAAACTGAAGCCTTCTCTCGG 1424
Db	
Qy	525 AACAAATTCACGGTCTTTGATGCTAGTCTAGACCGGAGCAGCAAGATGC-----A 575
Db	
Qy	1425 CACAAAGTTTACAATCTTCGACAGCCAGCCCTCCTCCTGACGCTGTAGTGTCTACCAACAA 1484
Db	
Qy	576 GAGAGCGGCTCTTTCAATTTTCAATGAATTTACCTAGAGTTCCTCAGGGAAGTTACCC 635
Db	
Qy	1485 TAAGACAAGCAAAAGGCAAGTCAAGCAAGTATCACCCGAGACTGCGGTAGGCAATTACAA 1544
Db	
Qy	636 CATCGCTCACATTTTCATACGAGTTTAAACGCTCTTAGGCTCTCGGCGACCGAGAAAGTGG 695
Db	
Qy	1545 TGTGTCTACAGTCAAGTACAGACTCACCGTCTTACCGCAA CAGGGGTCCAAGGAGATGCA 1604
Db	
Qy	696 TTGCATCATGGATACAATAC-----CTATGAGCATCTGGAGTCTCGGAGGAGTAG 745
Db	
Qy	1605 ATGCACCATGCATCTCAATACAGCTGAGTGTGATCTCCAGGAGGCGGGAAGGCCCTACCCC 1664
Db	
Qy	746 TAGGTTCAACATCGATAAGTCTTTTTCAGTGGTTCATCACCAGTCTTTAGTCTCTACT 805
Db	
Qy	1665 TACAGGCGCCATCCAGTCGCTTGACGAGCCAGGTGTGTGTCCACCTTACTTAGTGCCAAAGG 1724
Db	
Qy	806 CAAAACCATTCGCGCAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACTGGGAGATC 865
Db	
Qy	1725 GAAGGAAGTGGCTGTAGAAATCTCTTCGACAAAGCCTCAGCGCCGATCTGTCTGACCGCG 1784
Db	
Qy	866 -----CAACCATTTGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTT 911
Db	
Qy	1785 CTGGTGAACGGGAACACTCTAGTGTCTGAAGAACAAAGCTCTCTGTTGGATGAGCAGCT 1844
Db	

QY 912 ACCTTCTGGTGTAAATTTTCAATGTCGAGTCAAGTGGCTTCGGTTAAGAACTTTCA 971
DB 1845 GCAGTGTGGTGTCTCACTTCGGGGGGGTGTACCGTGTGCTGAGTCAAGTCTTCA 1904
QY 972 GCTTGGCAGTGTAGTGTGAAGCAGGGCAGACATCT-----GAGAGGATCAT 1022
DB 1905 GCTGTGCGCTCTGTGGTATCTTCCCTAAACATCCCGCAGCGGAGCAGGAGAGGTGAT 1964
QY 1023 ACTCCAGTTTGGAAAGTTGGNAGACATGTTTACCATGATATATGGATATCCGATTC 1082
DB 1965 CTTCCAGTTTGGAAAGATAGCAAGATATATTTACCATGGACTACCGGTACCCACTCTC 2024
QY 1083 TCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTGTGAAACAGAAATTCCTGTGAA 1140
DB 2025 GCGTTTCCAGGCTTTTGGCATCTGCTGACCGACTTCGACACCAACCGGCTTCGGA 2082

RESULT 14
US-10-767-701-12426
; Sequence 12426, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 12426
; LENGTH: 1183
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS9122_1
US-10-767-701-12426

Query Match 19.3%; Score 220.2; DB 7; Length 1183;
Best Local Similarity 62.3%; Pred. No. 1.1e-63;
Matches 454; Conservative 0; Mismatches 253; Indels 22; Gaps 6;

QY 428 GCGCAACTTGCATGTATATCATCTCTTTGGTTCAGACGATATCTCAAGAGAGCA 487
DB 4 GCGCCACATGCAAGACTAGTAAATTTCTTGTATAGGTTGATATGTCAAGGGAAGCA 63
QY 488 ACGCGTATCTTGGGAGATGAGATCGAATCTTCTTGGAAACAAATTCACGCTTTTGATG 547
DB 64 GCACCTATATTGGCAAGCTAAGATCAAACTTCTTGGAAACAAAGTTCACTGTCTATGATG 123
QY 548 GTAGTCAGAC---CGGAGCAGCGAAGATGCAGAAAGCCGCTCTTC-----TAATTT 595
DB 124 CTATCCACCATATGATGAGCTGTGTCTCAAGAGTGCCTGACCGTGTGTTGGTT 183
QY 596 TCATCAAGTTTCACTAGATGTTCTCAGGGAAGTTTACCCATCGCTCATATTTCTATAG 555
DB 184 TGAACCAAGGTCTCCCTTAGAGTTCCGGCTGGGAATTTATCTGTTTCAATATTTCTTAG 243
QY 656 AGTTAAACGTTTAGGCTCTCGGGACCGAGAAAGATGCTTCATCATGGATACAAATAC 715
DB 244 AGCTGAATGTTCTGGGCTCCAGAGGTCCAGAAAGGATGAATCTGTGTATGATTCATCC 303
QY 716 CTATGAGCATCTGTGGAGTCGCGAGGAGTAGTAGCTTCAACATCCATAAGCTCTTTTCCA 775
DB 304 CGGCATCAGCTGTTCGAGGAGGAGGAGAGCTCTACACAGACTGAATTTCCATTTAGCA 363
QY 776 GTCGGTCATCAGCATGCTTTAGGTCTCATCTCAAAACCATTTGGCGCATATATAGTGATCAT 835
DB 364 GTCTTGACTCTTTTCCCATCAATTTCCATTTCTTCAGATC-TAAATCAGCTCGGATAGACAGT 422
QY 836 GTAGCGACTCAGGCAACAACTCGGAGATCCACCATTTGGTGTGAGCAACAAAGCTCCAC 895

DB 423 TCAACATCGCAGTCACTCAGATCGGATA-----GGTTGGTGTGAAGCAAAAGTCTCCTTA 477
QY 896 GGTGCGATGACAGATTAC-GTTGCTGTGCTTAAATTTCCATGTCGAGTCACAGTGGCT 954
DB 478 GGTGCGATGAACAACCTGAGATGTTGGTGCCTGAAATTTCCGTGGACGGGTCACTGTGTCT 537
QY 955 TCGGTTAAGAACTTTTTCAGCTTGTGCGAGTTAGTGACTGTGAAGCAGGGCAGACATCTGAG 1014
DB 538 TCTGTTAAAACTTTTTCAGTTGGTGGCTTCTGATGACATGGACCAGGGAACCAAGATAAT 597
QY 1015 AGGA---TCATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTAAGA 1071
DB 598 AACAAAGTCACTCTCCAAATTTGGAAAGATTGGAAAGACTTGTTCACCATGACTACCGT 657
QY 1072 TATCCGATTTTCGCTTTCAAGCGTTTGTCTATCTGCTGAGCAGTGTTCGAAACCAAAAT 1131
DB 658 TATCCAATATCAGCATTTCAAGCTTTTGCATTTTGTCTGAGCAGTTTGTATACAAAAT 717
QY 1132 GCCTGTGAA 1140
DB 718 GCCTCGGA 726

RESULT 15
US-10-424-599-8380
; Sequence 8380, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 8380
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_107577C.1
US-10-424-599-8380

Query Match 19.3%; Score 219.8; DB 7; Length 2451;
Best Local Similarity 56.0%; Pred. No. 2.3e-63;
Matches 593; Conservative 0; Mismatches 382; Indels 84; Gaps 6;

QY 166 TGGCGCTGCGCGGAAACGTGTGGTGTGCGGGGTTTGTGTCGAGTGGAGGATCTC 225
DB 612 TGGCTGTCTCGAAAGCATGTTAGTGCATGTGTGTCAGTGTTCGAGTCTTGGAGGATATG 671
QY 226 ACCAAGGAGTTAGTGTCTCTGAAATTTCTCTTAAATTTGACTTTCCTTCTCTCCCTC 285
DB 672 TCAGAGCATGTTTAAAGCCAGAGTTTGTGGCAAACTTACATTCCTGTGTCTCTTG 731
QY 286 AAGCAGTCTGTGTCGAGGATTTCTCTAGTTCAATGCTTTTATAAAACGTAATCGAAATAC 345
DB 732 AAGCAGCTTGGCCACGGATGGAATCATTCATGTTTATCAAAAGAGATAAATCTAAT 791
QY 346 CAATGATCATCTCTATCTCGGATTA---CTACCTCTTTGACGAGTAACCGGAGTTT 402
DB 792 TTAACATACCACTATTCCTTTGTCTCAGCCCTGCTTTGTTAGTTGAAAATGGAAAATTC 851
QY 403 CTTCTTGTGCTTCTTAAGCTGAAGCGCAACTTGCACTGATTACATCATCTCTTTGCGT 462
DB 852 CTCCTTTCTGCTAAGAGGCAAGAGAGAACACTTACAGAGTACGTTATTTCCATGGAT 911
QY 463 TCAGACGATATCTCAAGAGAAAGCAACGCGTATCTTTGGAGAAATGAGATCGAACTTCCT 522

```
Db 912 GCTGACACATCTCTAGATCCAGTAACATTTACATTTGGAAGCTGAGATCAAAATTTTCTT 971
Qy 523 GGAACAAAATTCACGGT-----CTTTGATGGTAGTCAGACGGGA 561
Db 972 GGTACAAAGATTCAATTATATGCGATACACAGCCCTCCATACCTCTTCTGCCCATATATGCCCT 1031
Qy 562 GCAGCGAAGATGCAGAGAGCGCTCTTCTAAATTCATCAAGTTTCACCTAGAGTTCTT 621
Db 1032 CCTATGACTGGGAAGACGAGCCGTAGATTTATTCAAAAGGTCTCGCCTAAGGTCCCA 1091
Qy 622 CAGGGAAGTTACCCCATCGCTCAATTTCATACGAGTTAAACGCTTTAGGCTCTCGGGGA 681
Db 1092 TCTGGGAGTTACAACTAGCTCAGGTAAATATGATGAATTAATGTGCTTGGAACTCGAGGC 1151
Qy 682 CCGAGAAGAAATGCGTTGCTCATGGAATCAATACC-----TATGAGCA 724
Db 1152 CCAAGAAAGATGCACTGCGTTATGCAATCAATACCAAGCTTCAGCACTTGATGCGGCGGC 1211
Qy 725 TCGTGGAGTCGGGAGGAGTAGTAGTTCACATCCATAG----CTCTTTTCCAGTCGG 780
Db 1212 ACTGCTCTGGCCAGCCAGAGCTTCTCTGTCCTTGGAGACTCGTTTCGGAGCATC 1271
Qy 781 TCATCACCAAGTCCTTTAGGTCTCACTCAAAACCAATTCGCGAGTAATAGTGCATCATGTAGC 840
Db 1272 TCGTTTTCAAAGTCTCTAGATCGTTCGATGAGTTCAGCAGTTCACGATTTTCTGAGATT 1331
Qy 841 GACTCAGGCAACAACTCGGAGATC-----CACCATTGGTGTGAGCAACAAA 888
Db 1332 GGGGAATCTATCATCGAGGATGATGATGGCAAGATGAGACCCCTTGGTTCTGAAAAACAAG 1391
Qy 889 GCTCCAGGTTGGCATGAGCAGATTACGTTGCTGGTGTCTAAATTTCCATGTCGAGTCACA 948
Db 1392 CCTCCAAGATGGCAAGCAGTTACAAATGTGGTGCCTTAATTTCCGTGGACGAGTAACA 1451
Qy 949 GTGGCTTCGGTTAAGAACTTTACAGCTTGTGGCAGTTAGTGAAGCAGCGGCAGACA 1008
Db 1452 GTTGGCTCTGTTAAAAAATTTCAAGTTGATGTGTCGCCACCACCCGCTGGTGCACT 1511
Qy 1009 TC-----TGAGAGGATCATCTCCAGTTTGGGAAAGTT 1041
Db 1512 ACACCATCTCAACCAAGCTCCACCGGAGCATGATAAAATCATCTTCAATTTGGCAAAGTT 1571
Qy 1042 GGGAGGACATGTTTACCATGGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGCT 1101
Db 1572 GGTAAAGACATGTTTACCATGGATTATCCGTACCCCTTATCTGCAATTCGAAGCTTTTGA 1631
Qy 1102 ATCTGCTGAGCAGTTTGAACCAAGAAATTCCTGTGAA 1140
Db 1632 ATATGCTTGAGCAGCTTTGACACCAAAATTTGGCTTTGAA 1670
```

Search completed: December 26, 2005, 07:12:36
Job time : 1110 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model :
Run on: December 26, 2005, 01:49:06 ; Search time 5217 Seconds
(without alignments)
10223.744 Million cell updates/sec

Title: US-10-763-042-20
Perfect score: 1140
Sequence: 1 agaggttcgaagttact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1140	100.0	1560	4	CNS0A4PF
2	1127.2	98.9	1261	4	US0A5VU
3	1090.8	95.7	1438	4	CNS0A5QU
4	1004.8	88.1	1535	4	CNS0A5IP
5	830.2	72.8	1036	5	US0A5835
6	598.2	52.5	795	5	US0A3792
7	592.8	52.0	597	2	BE523610
8	569.2	49.9	1281	4	CNS0A1DW
9	516	45.3	516	1	AV527654
10	466	40.9	637	5	BU636137
11	449	39.4	598	1	AU237181
12	422.6	37.1	704	5	BU635249
13	398	34.0	420	8	F13831
14	333	29.2	728	5	BU635253
15	331.2	29.1	358	8	T44322
16	319.6	28.0	982	7	CK277179
17	308	27.0	934	7	CV267012
18	300.8	26.4	650	6	CD819805
19	298.4	26.2	391	1	AI993972
20	292	25.6	509	1	AV521502
21	285.8	25.1	909	8	CX546892
22	282.4	24.8	904	8	CX667858

23	270.8	23.8	351	8	T44750
24	259.4	22.8	577	8	DN776642
25	259	22.7	694	8	CX297509
26	258	22.6	811	8	CX673843
27	254.2	22.3	667	5	BU890728
28	253.8	22.3	876	7	CV260993
29	253.6	22.2	837	7	CK265631
30	250.4	22.0	632	5	BU827262
31	248.8	21.8	614	2	BE021478
32	245.8	21.6	663	2	BF004440
33	245.8	21.6	666	6	CAB00472
34	244.8	21.5	933	8	CX667859
35	240.8	21.1	779	7	CV710051
36	240.8	21.1	809	6	CD576263
37	240.2	21.1	578	8	CX195446
38	239	21.0	344	6	CB977099
39	235.6	20.7	739	5	BU834291
40	235	20.6	548	1	AU235780
41	232.8	20.4	726	7	CO168454
42	231	20.3	232	9	BZ767812
43	228.8	20.1	713	8	DT018763
44	228.6	20.1	897	8	DR540496
45	227.4	19.9	1107	7	CK209327

ALIGNMENTS

RESULT 1
CNS0A4PF
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

CNS0A4PF 1560 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLT7B4ZA09 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
BX822936
HTC; GSLT cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1560)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen)
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
Location/Qualifiers
1. .1560
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

		/db_xref="taxon:3702" /clone="GSLTFB74ZA09" /tissue_type="flowers and buds" /ecotype="Col-0" /plasmid="pCWSPORT_6" 1. .1560 /gene="At3g06380"	
gene			
ORIGIN			
Query Match		100.0%; Score 1140; DB 4; Length 1560;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1140; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGACGTCCTCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTCTAGTTCACGCC	60
Db	261	ATGACGTCCTCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTCTAGTTCACGCC	320
QY	61	GCGCCTCAACCGCTAATAGTTTCAGACCCCTTTTCAGCTGGTCGGAGCTCCCGAGGAGCTG	120
Db	321	GCGCCTCAACCGCTAATAGTTTCAGACCCCTTTTCAGCTGGTCGGAGCTCCCGAGGAGCTG	380
QY	121	CTTAGAGAAATCTGATTAGGTTGAGACTTTGACGGCGGCGATTGGCGTTCGGCGCA	180
Db	381	CTTAGAGAAATCTGATTAGGTTGAGACTTTGACGGCGGCGATTGGCGTTCGGCGCA	440
QY	181	AACTGCTGGCTTGTCGCGCGTTCCTGCTAGCTGGAGGATTTCTACCAAGGAGATTGTA	240
Db	441	AACTGCTGGCTTGTCGCGCGTTCCTGCTAGCTGGAGGATTTCTACCAAGGAGATTGTA	500
QY	241	GCTGTTCTCGAATTCCTCTAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGTGCA	300
Db	501	GCTGTTCTCGAATTCCTCTAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGTGCA	560
QY	301	AGAGATTCTAGTTCAATGCTTTATAAAGTAATCGAATACTCAATCGTATCATCTC	360
Db	561	AGAGATTCTAGTTCAATGCTTTATAAAGTAATCGAATACTCAATCGTATCATCTC	620
QY	361	TATCTCGGATTAATCACTCTTTGCGGATACGGGAAGTTTCTTCTGCTCTCTAAG	420
Db	621	TATCTCGGATTAATCACTCTTTGCGGATACGGGAAGTTTCTTCTGCTCTCTAAG	680
QY	421	CTGAACGGCGCACTTGCACTGATTAATCATCATCTCTTTGCGTTTCAGACGATATCTCAAG	480
Db	681	CTGAACGGCGCACTTGCACTGATTAATCATCATCTCTTTGCGTTTCAGACGATATCTCAAG	740
QY	481	AGAAGCAACCGGATCTTGGGAGAAATGAGATCGAACTTCTCTTGGAAACAAAATTCACGGTC	540
Db	741	AGAAGCAACCGGATCTTGGGAGAAATGAGATCGAACTTCTCTTGGAAACAAAATTCACGGTC	800
QY	541	TTTGATGGTAGTCAGACCGGAGCAGGAGATGCGAAGAGCGGCTCTTCTAATTTCTATC	600
Db	801	TTTGATGGTAGTCAGACCGGAGCAGGAGATGCGAAGAGCGGCTCTTCTAATTTCTATC	860
QY	601	AAAGTTTCACTAGAGTTCTCAGGGAAGTTTACCCATCGCTCACAATTTTCATACAGATTA	660
Db	861	AAAGTTTCACTAGAGTTCTCAGGGAAGTTTACCCATCGCTCACAATTTTCATACAGATTA	920
QY	661	AACGTTTAGGCTCTCGGGACCGAGAGAATCGTTGCAATCATGATGATACAATCTATG	720
Db	921	AACGTTTAGGCTCTCGGGACCGAGAGAATCGTTGCAATCATGATGATACAATCTATG	980
QY	721	AGCATCTGGAGTCGGAGGAGTAGTCTTCAACATCCATAGCTCTTTTCCAGTCGG	780
Db	981	AGCATCTGGAGTCGGAGGAGTAGTCTTCAACATCCATAGCTCTTTTCCAGTCGG	1040
QY	781	TCATCACCACTCTTTAGGCTCTCACTCAAAACCACTGGCAGTAATAGTGCATCATGTAGC	840
Db	1041	TCATCACCACTCTTTAGGCTCTCACTCAAAACCACTGGCAGTAATAGTGCATCATGTAGC	1100
QY	841	GACTCAGGCAACCACTGGGAGATCCACCAATGGTCTGAGCAACAAAGCTCCACGGTGG	900
Db	1101	GACTCAGGCAACCACTGGGAGATCCACCAATGGTCTGAGCAACAAAGCTCCACGGTGG	1160
QY	901	CATGAGCAGTTACGTTGCTGGTCTTAAATTTCCATGGTCGAGTCACAGTGCCTTCGGTT	960
Db	1161	CATGAGCAGTTACGTTGCTGGTCTTAAATTTCCATGGTCGAGTCACAGTGCCTTCGGTT	1220
QY	961	AAGAACTTTACGTTGCTGGCAGTTAGTCACTGTGAAGCAGGCGACACATCTCAGAGGATC	1020
Db	1221	AAGAACTTTACGTTGCTGGCAGTTAGTCACTGTGAAGCAGGCGACACATCTCAGAGGATC	1280
QY	1021	ATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTTATCGATATCCGATT	1080
Db	1281	ATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTTATCGATATCCGATT	1340
QY	1081	TCGCGTTTCAAGCGTTTGCCTATCTGCTGAGCAGTTTGAACACGAGATTCCTCTGTGAA	1140
Db	1341	TCGCGTTTCAAGCGTTTGCCTATCTGCTGAGCAGTTTGAACACGAGATTCCTCTGTGAA	1400
RESULT 2			
CNS0A5VU		1261 bp mRNA linear HTC 06-FEB-2004	
LOCUS		Arabidopsis thaliana Full-length cDNA Complete sequence from clone	
DEFINITION		GSLTFB1ZD02 of Flowers and buds of strain col-0 of Arabidopsis	
ACCESSION		thaliana (thale cress).	
VERSION		BX822204	
KEYWORDS		HTC; GSLT cDNA.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;	
TITLE		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
JOURNAL		1 (bases 1 to 1261)	
REFERENCE		Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,	
AUTHORS		Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,	
TITLE		Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.	
JOURNAL		Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:	
COMMENT		A Combined Approach to Evaluate and Improve Arabidopsis Genome	
JOURNAL		Annotation	
REFERENCE		Unpublished	
AUTHORS		2 (bases 1 to 1261)	
TITLE		Genoscope.	
JOURNAL		Direct Submission	
COMMENT		Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :	
JOURNAL		BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
COMMENT		- Web : www.genoscope.cns.fr)	
JOURNAL		The sequences are based on single pass reads.	
REFERENCE		Life Technologies (a division of Invitrogen)	
AUTHORS		full-length libraries construction : Temple G.	
TITLE		Genoscope members carried out sequencing and annotation : Castelli	
JOURNAL		V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,	
COMMENT		Schachter V., Weissenbach J., Salanoubat M.	
JOURNAL		URGV INRA : Clepet C., Caboche M.	
REFERENCE		Annotation is based on the June 2003 version of the Arabidopsis	
AUTHORS		genome released by MIPS (Munich Information center for Protein	
TITLE		Sequences). 5 prime and 3 prime are assembled with Phrap.	
JOURNAL		http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full	
COMMENT		length	
JOURNAL		http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.	
REFERENCE		Location/Qualifiers	
AUTHORS		1. .1261	
TITLE		/organism="Arabidopsis thaliana"	
JOURNAL		/mol_type="mRNA"	
COMMENT		/db_xref="taxon:3702"	
JOURNAL		/clone="GSLTFB1ZD02"	
REFERENCE		/tissue_type="Flowers and buds"	
AUTHORS		/ecotype="Col-0"	
TITLE		/plasmid="pCWSPORT_6"	
JOURNAL		1. .1261	
COMMENT		/gene="At3g06380"	
JOURNAL		gene	
REFERENCE		ORIGIN	
AUTHORS		Query Match	
TITLE		Best Local Similarity	
JOURNAL		98.9%; Score 1127.2; DB 4; Length 1261;	
COMMENT		99.3%; Pred. No. 0;	

Matches 1132; Conservative 0; Mismatches 8; Indels 0; Gaps 0;									
QY	1	ATGACGTTCCGAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTCAAGCC	60						
DB	57	ATGACGTTCCGAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTCAAGCC	116						
QY	61	GCGGCTCAACCGCTAATAGTTTCAGACCCCTTTTCAGCTGTCGAGCTCCGCGAGAGCTG	120						
DB	117	GCGGCTCAACCGCTAATAGTTTCAGACCCCTTTTCAGCTGTCGAGCTCCGCGAGAGCTG	176						
QY	121	CTTAGAGAAATCTGATTAGGTTTGAGACTGTTTGACGGCGGCGATTGGCGCTCGCGCGA	180						
DB	177	CTTAGAGAAATCTGATTAGGTTTGAGACTGTTTGACGGCGGCGATTGGCGCTCGCGCGA	236						
QY	181	AACGTGCTGCTGTCGCGCGTTTCTGCTAGCTGAGGATTCTCACGAAGGAGATTGA	240						
DB	237	AACGTGCTGCTGTCGCGCGTTTCTGCTAGCTGAGGATTCTCACGAAGGAGATTGA	296						
QY	241	GCTGTTCTCGAATCTCCTCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGCTCA	300						
DB	297	GCTGTTCTCGAATCTCCTCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGCTCA	356						
QY	301	AGAGATTCTCTAGTTCAATGCTTTTATAAACGTAATCGAAATACTCAATCGTATCATCTC	360						
DB	357	AGAGATTCTCTAGTTCAATGCTTTTATAAACGTAATCGAAATACTCAATCGTATCATCTC	416						
QY	361	TATCTCGGATTAACTACTCTTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTCTAAG	420						
DB	417	TATCTCGGATTAACTACTCTTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTCTAAG	476						
QY	421	CTGAGCGCGCACTGCACTGATTACATCATCTCTTTCGCTTCAGCATATCTCAAG	480						
DB	477	CTGAGCGCGCACTGCACTGATTACATCATCTCTTTCGCTTCAGCATATCTCAAG	536						
QY	481	AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTGGAAACAAATTCAGGTC	540						
DB	537	AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTGGAAACAAATTCAGGTC	596						
QY	541	TTTGATGTTAGTCAGACCGGAGCGAAGATGACGAAGAGCGGCTTCTTAATTTATC	600						
DB	597	TTTGATGTTAGTCAGACCGGAGCGAAGATGACGAAGAGCGGCTTCTTAATTTATC	656						
QY	601	AAAGTTTACCTAGATTCTCAGGAAAGTTACCCATCGCTCACATTTTACACGATT	660						
DB	657	AAAGTTTACCTAGATTCTCAGGAAAGTTACCCATCGCTCACATTTTACACGATT	716						
QY	661	AACTCTTAGGCTCTCGGAGCGGACGAAGAATGCGTTGTCATCATGATACATACCTATG	720						
DB	717	AACTCTTAGGCTCTCGGAGCGGACGAAGAATGCGTTGTCATCATGATACATACCTATG	776						
QY	721	AGCATCGTGGAGTCGCGAGGAGTAGTCTTCAACATCCATAGCTCTTTTCCAGTCGG	780						
DB	777	AGCATCGTGGAGTCGCGAGGAGTAGTCTTCAACATCCATAGCTCTTTTCCAGTCGG	836						
QY	781	TCATCACGACTTTTAGTCTCCTCAAAACCAATTCGGCAGTAATAGTGCATCATGTAGC	840						
DB	837	TCATCACGACTTTTAGTCTCCTCAAAACCAATTCGGCAGTAATAGTGCATCATGTAGC	896						
QY	841	GACTCAGGCAACAACCTGGGAGATCCACCATTTGGTCTGAGCAACAAGCTCCAGCTGG	900						
DB	897	GACTCAGGCAACAACCTGGGAGATCCACCATTTGGTCTGAGCAACAAGCTCCAGCTGG	956						
QY	901	CATGACGATTTAGTGTGCTGCTTAAATTTCCATGGTTCGAGTCACAGTGGCTTCGGTT	960						
DB	957	CATGACGATTTAGTGTGCTGCTTAAATTTCCATGGTTCGAGTCACAGTGGCTTCGGTT	1016						
QY	961	AAGAACTTTTTCAGTTTGGCAGTTTGTGTAAGCAAGGAGGAGATCTGAGAGGATC	1020						
DB	1017	AAGAACTTTTTCAGTTTGGCAGTTTGTGTAAGCAAGGAGGAGATCTGAGAGGATC	1076						
QY	1021	ATACTCCAGTTTGGGAAAGTTGGGAGGACATGTTTACCATGGATTTATCGGATTCGGAT	1080						
DB	1077	ATACTCCAGTTTGGGAAAGTTGGGAGGACATGTTTACCATGGATTTATCGGATTCGGAT	1136						

QY	1081	TCCTGCTTCAAGCGTTTCTCTATCTGCTGAGCAGTTTGAACACGAAATTCCTGTGAA	1140						
DB	1137	TCCTGCTTCAAGCGTTTCTCTATCTGCTGAGCAGTTTGAACACGAAATTCCTGTGAA	1196						
RESULT 3									
CNS0A5QU									
LOCUS									
DEFINITION									
Arabidopsis thaliana Full-length cDNA complete sequence from clone									
GSLTGH27D03 of Hormone Treated Callus of strain col-0 of									
Arabidopsis thaliana (thale cress).									
ACCESSION									
BX824160.1 GI:42464062									
VERSION									
HTC; GSLT_cDNA.									
KEYWORDS									
Arabidopsis thaliana (thale cress)									
SOURCE									
ORGANISM									
Arabidopsis thaliana									
REFERENCE									
AUTHORS									
Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C.,									
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,									
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.									
TITLE									
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:									
A Combined Approach to Evaluate and Improve Arabidopsis Genome									
Annotation									
Unpublished									
REFERENCE									
2 (bases 1 to 1438)									
Genoscope.									
Direct Submission									
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :									
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr									
- Web : www.genoscope.cns.fr)									
COMMENT									
The sequences are based on single pass reads.									
Life Technologies (a division of Invitrogen) members carried out									
full-length libraries construction : Temple G.									
Genoscope members carried out sequencing and annotation : Castelli									
V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,									
Schachter V., Weissenbach J., Salanoubat M.									
URGV INRA : Clepet C., Caboche M.									
Annotation is based on the June 2003 version of the Arabidopsis									
genome released by MIPS (Munich Information center for Protein									
Sequences). 5 prime and 3 prime are assembled with Phrap.									
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_BP/Full									
length									
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.									
FEATURES									
source									
1..1438									
Location/Qualifiers									
/organism="Arabidopsis thaliana"									
/mol_type="mRNA"									
/db_xref="taxon:3702"									
/clone="GSLTGH27D03"									
/tissue_type="Hormone Treated Callus"									
/ecotype="Col-0"									
/plasmid="pCMVSPORT_6"									
1..1438									
/gene="At3g06380"									
ORIGIN									
Query Match									
Best Local Similarity 95.7%; Score 1090.8; DB 4; Length 1438;									
Matches 1136; Conservative 0; Mismatches 2; Indels 4; Gaps 4;									
QY	1	ATGAGCTTCCGAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTCAAGCC	60						
DB	198	ATGAGCTTCCGAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTCAAGCC	257						
QY	61	GCGGCTCAACCGCTAATAGTTTCAGACCCCTTTTCAGCTGTCGAGCTCCGCGAGAGCTG	120						
DB	258	GCGGCTCAACCGCTAATAGTTTCAGACCCCTTTTCAGCTGTCGAGCTCCGCGAGAGCTG	317						
QY	121	CTTAGAGAAATCTCTGATTAGGTTTGAGACTGTTTGACGGCGGCGATTGGCGCTCGCGCGA	180						

Db	318		CTTAGAGAAATCCTGATTAGGTTGAGACTGTTGACGGCGCGGATTTGGCCGTCGGCGA	377
Qy	181		AACGTGGTGGCTTGTGCGCGGCTTTGTCTAGCTGGAGGATTCACCAAGAGATTGTA	240
Db	378		AACGTGGTGGCTTGTGCGCGGCTTTGTCTAGCTGGAGGATTCACCAAGAGATTGTA	437
Qy	241		GCTGTTCTCGAATCTTCCTCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCA	300
Db	438		GCTGTTCTCGAATCTTCCTCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCA	497
Qy	301		AGAGATTCTCTAGTTCAATCTTTATAAACGTAATCGAAATFACTCAATCGTATCATCTC	360
Db	498		AGAGATTCTCTAGTTCAATCTTTATAAACGTAATCGAAATFACTCAATCGTATCATCTC	557
Qy	361		TATCTCGGATTAATACCTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTCTTAAG	420
Db	558		TATCTCGGATTAATACCTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTCTTAAG	617
Qy	421		CTGAAGCGCGCAACTTGCACTGATTACATCATCTCTTTGCGTTTCAGACGATATCTCAAG	480
Db	618		CTGAAGCGCGCAACTTGCACTGATTACATCATCTCTTTGCGTTTCAGACGATATCTCAAG	677
Qy	481		AGAAGCAACCGCTATCTTGGGAGATGAGATCGAACTTCTTTGGAAACAAATTCACGGTC	540
Db	678		AGAAGCAACCGCTATCTTGGGAGATGAGATCGAACTTCTTTGGAAACAAATTCACGGTC	737
Qy	541		TTTGATGTAGTCAGACCGGAGCAGGAAGATGCGAGAGCGGCTCTTCTAATTTTCATC	600
Db	738		TTTGATGTAGTCAGACCGGAGCAGGAAGATGCGAGAGCGGCTCTTCTAATTTTCATC	797
Qy	601		AAAGTTTACCTAGAGTTCTCTCAGGGAAGTTACCCATCGCTCACATTTTCATACAGTTA	660
Db	798		AAAGTTTACCTAGAGTTCTCTCAGGGAAGTTACCCATCGCTCACATTTTCATACAGTTA	857
Qy	661		AACGCTTAGGCTCTCGGG-ACCAGAGAAGATGCGTTGCA-TCATGGATACAAATACCTTA	718
Db	858		AACGCTTAGGCTCTCGGGAAACCGAGAGAATGCGTTGCAATCATGGATACAAATACCTTA	917
Qy	719		TGAGCATCTGGAGTCGCGAGAGTAGTAGCTTCAACATCCATAGCTCTTTTTCAGTTC	778
Db	918		TGAGCATCTGGAGTCGCGAGAGTAGTAGCTTCAACATCCATAGCTC-TTTTTCAGTTC	976
Qy	779		GCTCATCACCAGCTTTTAGGTTCTCACTCAAACCAATGCGCAGTAAATAGTCATCATGTA	838
Db	977		GCTCATCACCAGCTTTTAGGTTCTAACTCAAACCAATGCGCAGTAAATAGTCATCATGTA	1036
Qy	839		GCGACTCAGGCAACAACTGGGAGATCCACCATTGGTGTGCTGAGCAACAAAGCTCCACGCT	898
Db	1037		GCGACTCAGGCAACAACTGGGAGATCCACCATTGGTGTGCTGAGCAACAAAGCTCCACGCT	1096
Qy	899		GGCATGAGCAGTTACGTTGCTGGTCTTAAATTTTCCATGGTGAAGTCAGTGGCTTCGG	958
Db	1097		GGCATGAGCAGTTACGTTGCTGGTCTTAAATTTTCCATGGTGAAGTCAGTGGCTTCGG	1156
Qy	959		TTAAGAACTTTTCAAGCTTTGGCAGTTAGTAGCTGTGAAGCAGGCGAGACATCTGAGAGGA	1018
Db	1157		TTAAGAACTTTTCAAGCTTTGGCAGTTAGTAGCTGTGAAGCAGGCGAGACATCTGAGAGGA	1216
Qy	1019		TCATATCTCCAGTTTGGGAAGTTGGGAAGACATGTTTACCATGGATTATGGATATCCGA	1078
Db	1217		TCATATCTCCAGTTTGGGAAGTTGGGAAGACATGTTTACCATGGATTATGGATATCCGA	1276
Qy	1079		TTTCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGGTTTGAACACAGAAATTCGCTGTG	1138
Db	1277		TTTCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGGTTTGAACACAGAAATTCGCTGTG	1335
Qy	1139	AA	1140	
Db	1336	AA	1337	

CNS0A51P	1535 bp	mRNA	linear	HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone			
DEFINITION	GSUTPGH44ZH11 of Hormone Treated Callus of strain col-0 of			
	Arabidopsis thaliana (Chale cress).			
ACCESSION	BX824384			
VERSION	1			
KEYWORDS	HTC; GSLT cDNA.			
SOURCE	Arabidopsis thaliana (Chale cress)			
ORGANISM	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;			
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE	1 (bases 1 to 1535)			
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,			
	Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,			
	Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.			
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:			
	A Combined Approach to Evaluate and Improve Arabidopsis Genome			
JOURNAL	Annotation			
AUTHORS	Unpublished			
REFERENCE	2 (bases 1 to 1535)			
TITLE	Direct Submission			
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :			
	BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr			
	- Web : www.genoscope.cns.fr)			
COMMENT	The sequences are based on single pass reads.			
	Life Technologies (a division of Invitrogen) members carried out			
	full-length libraries construction : Temple G.			
	Genoscope members carried out sequencing and annotation : Castelli			
	V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,			
	Schachter V., Weissenbach J., Salanoubat M.			
	URGV INRA : Clepet C., Caboche M.			
	Annotation is based on the June 2003 version of the Arabidopsis			
	genome released by MIPS (Munich Information center for Protein			
	Sequences) . 5 prime and 3 prime are assembled with Phrap.			
	http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full			
	length			
FEATURES	Location/Qualifiers			
source	1..1535			
	/organism="Arabidopsis thaliana"			
	/mol_type="mRNA"			
	/db_xref="taxon:3702"			
	/clone="GSUTPGH44ZH11"			
	/tissue_type="Hormone Treated Callus"			
	/ecotype="Col-0"			
gene	1..1535			
	/plasmid="pCMVSPORT_6"			
	/gene="At3g06380"			
ORIGIN				
Query Match	88.1%; Score 1004.8; DB 4; Length 1535;			
Best Local Similarity	94.1%; Pred. No. 1.2e-305;			
Matches 1078;	Conservative 0; Mismatches 62; Indels 6; Gaps 3;			
Qy	1 ATGACGTTCCGAGTTTACTCCAGGAATCGGTTAGCCACACCGTGTAGTTCACGCC	60		
Db	200 ATGACGTTCCGAGTTTACTCCAGGAGATGCGGTTAGGCCACACCGTGTAGTTCACGCC	259		
Qy	61 GCCGCTCAACCGCTTAATAGTTTCAGCTGTCGAGCTCCCGGAGGAGCTG	120		
Db	260 GCCGCTCAACCGCTTAATAGTTTCAGCTGTCGAGCTCCCGGAGGAGCTG	319		
Qy	121 CTTAGAGAAATCCTGATTAGGGTTGAGACTGTTGACGCGCGGATTTGGC-CGTGCGGG	179		
Db	320 CTTAGAGAGATGCTGATTAGGGTTGAGACTGTTGACGCGCGGATTTGGC-CGTGCGGG	379		
Qy	180 AAACGTGGTGGCTTGTGCGCGGCTTTGTCGTAGCTGGAGGATTTCTCACCAGGAGATTGT	239		
Db	380 AAACGTGGTGGCTTGTGCGCGGCTTTGTCGTAGCTGGAGGATTTCTCACCAGGAGATTGT	439		


```

QY 240 AGCTGTTCCGTAATCTCCCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCC 299
Db 440 AGCTGTTCCGTAATCTCCCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCC 499
QY 300 AAGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAATCGAAATACCAATCGTATCATCT 359
Db 500 AAGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAATCGAAATACCAATCGTATCATCT 559
QY 360 CTATCTCGGATTAACCTCTTTGAGGATTAACGGGAAGTTT-CTTCTGCTGCTTCTA 418
Db 560 CTATCTCGGATTAACCTCTTTGAGGATTAACGGGAAGTTTCTTCTGCTGCTTCTG 619
QY 419 AGCTGAAGCGCGCAACTTGCACGTATTACATCATCTCTTTGGCTTCAGACGATATCTCAA 478
Db 620 AGCTGAAGCGCGCGAGTTGCATCATCTATCATCTGCTTTGCTTCAGACGATATCGCAA 679
QY 479 AGAAGAGCAACCGCTATCTTTGGAGATGAGATCGAACTTCTTTGGAACAAAATTCACGG 538
Db 680 AGAAGAGCAACCGCTATGTTGGGAGATGAGATCGAACTTCTTTGGAACAAAATTCACGG 739
QY 539 TCTTTGATGTTAGTACAGCGGAGCGAAGATGAGAGAGCCGCTCTTCTTAATTTCA 598
Db 740 TCTTTGATGTTAGTACAGCGGAGCGAAGATGAGAGAGCCGCTCTTCTTAATTTCA 799
QY 599 TCAAGTTTCACTAGATTCTTCCAGGAGTTTACCCTATGCTCACATTTTCATACGAGT 658
Db 800 TCAAGTTTCACTAGATTCTTCCAGGAGTTTACCCTATGCTCACATTTTCATACGAGT 859
QY 659 TAAACGCTTTAGGCTCTCGGGACCGAGAGAAATGCTTGTCATCATGATACATACTCA 718
Db 860 TAAACGCTTTAGGCTCTCGGGACCGAGAGAAATGCTTGTCATCATGATACATACTCA 919
QY 719 TGAGCATCTGAGTCCGAGGATGATGATCTTCAACATCCATAAGCTCTTTTTC- 774
Db 920 TGAGCATCTGAGTCCGAGGATGATGATCTTCAACATCCATAAGCTCTTTCGCGAGT 979
QY 775 AGTCCGTCATACCAGTCTTTAGTCTCACTCAAAACCAATGCGCAGTAAATAGTCATCA 834
Db 980 CGGTCATCACGAGTCTTTGCTCTCACTCAAAACCAATGCGCAGTAAATAGTCATCA 1039
QY 835 TGAGCGACTCAGGCAACCACTGGGAGATCCACCAATGCTGTGAGCAACAAAGTCCA 894
Db 1040 TGAGCGACTCAGGCAACCACTGGGAGATCCACCAATGCTGTGAGCAACAAAGTCCA 1099
QY 895 CGGTGCGATGAGCAGTTACGTTGCTGCTTAAATTTCCATGTCAGTCAAGTGGCT 954
Db 1100 CGGTGCGATGAGCAGTTACGTTGCTGCTTAAATTTCCATGTCAGTCAAGTGGCT 1159
QY 955 TCGGTTAAGAACTTTGAGTCTGAGTCTGAGTCTGAGCAGGCGAGATCTGAG 1014
Db 1160 TCGGTTAAGAACTTTGAGTCTGAGTCTGAGTCTGAGCAGGCGAGATCTGAG 1219
QY 1015 AGGATCATCTCCAGTTTGGGAAAGTTGGGAGGACATGTTTACCATGATATGATAT 1074
Db 1220 AGGATCATCTCCAGTTTGGGAAAGTTGGGAGGACATCTTACCATGATATGATAT 1279
QY 1075 CCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTTGGAAACCGAATTTGCC 1134
Db 1280 CCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTCTTGGAAACCGAATTTGCC 1339
QY 1135 TGTGAA 1140
Db 1340 TGTGAA 1345

RESULT 5
BX838835 1036 bp mRNA linear EST 11-FEB-2004
LOCUS
DEFINITION
BX838835 Arabidopsis thaliana Flowers and buds Col-0 Arabidopsis
thaliana cDNA clone GSLTFFB38ZC05 SPRIM, mRNA sequence.
ACCESSION
BX838835
VERSION
BX838835.1 GI:42532918
KEYWORDS
EST.

```

```

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsia.
REFERENCE 1 (bases 1 to 1036)
AUTHORS Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished (2004)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_BP/EST
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
1..1036
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="GSLTFFB38ZC05"
/tissue_type="Flowers and buds"
/clone_lib="Arabidopsis thaliana Flowers and buds Col-0"
ORIGIN
Query Match 72.8%; Score 830.2; DB 5; Length 1036;
Best Local Similarity 99.4%; Pred. No. 1.8e-250;
Matches 854; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 1 ATGAGCTTCCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTGTTACGCC 60
Db 173 ATGAGCTTCCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTGTTACGCC 232
QY 61 GCGGCTCAACCGCTTAATAGTTTCAACCCCTTTTCACTGCTGCGAGCTCCGGAGGAGCTG 120
Db 233 GCGGCTCAACCGCTTAATAGTTTCAACCCCTTTTCACTGCTGCGAGCTCCGGAGGAGCTG 292
QY 121 CTTAGAGAAATCTGATTAGGGTTGAGACTGTTGACGGCGGATTTGGCGCTCGCGCGCA 180
Db 293 CTTAGAGAAATCTGATTAGGGTTGAGACTGTTGACGGCGGATTTGGCGCTCGCGCGCA 352
QY 181 AACGTGTGGCTTGTGCGCGCGTTTGTCTGAGCTGGAGATTCTCACCAAGGAGATTGTA 240
Db 353 AACGTGTGGCTTGTGCGCGCGTTTGTCTGAGCTGGAGATTCTCACCAAGGAGATTGTA 412
QY 241 GCTGTTCTCGAATTTCTCTTAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGTGCA 300
Db 413 GCTGTTCTCGAATTTCTCTCTTAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGTGCA 472
QY 301 AGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAATCGAAATACCAATCGTATCATCTC 360
Db 473 AGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAATCGAAATACCAATCGTATCATCTC 532
QY 361 TATCTCGGATTAACCTCTTTTGGAGGATAACGGGAAGTTTCTTCTGCTGCTTCTTAAG 420
Db 533 TATCTCGGATTAACCTCTTTTGGAGGATAACGGGAAGTTTCTTCTGCTGCTTCTTAAG 592
QY 421 CTGAAGCGCGCAACTTGGCACTGATTACATCTCTTTTGGTTCAGACGATATCTCAAAG 480

```

|||||
Db 593 CTGAAGCGGCAACTGCTGACTGATTACATCATCTCTTTGCGTTCAGAGATATCTCAAG 652
QY 481 AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTCTCTTGGAAACAAAATTCACGGTC 540
Db 653 AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTCTCTTGGAAACAAAATTCACGGTC 712
QY 541 TTTGATGTTAGTTCAGACCGGAGCGGAGATGCGAGAGCGGCTCTCTTAATTTTCATC 600
Db 713 TTTGATGTTAGTTCAGACCGGAGCGGAGATGCGAGAGCGGCTCTCTTAATTTTCATC 772
QY 601 AAAGTTTACCTAGAGTTCTCTCAGGGAAGTTTACCCCATCGCTCACATTTTCATACGAGTTA 660
Db 773 AAAGTTTACCTAGAGTTCTCTCAGGGAAGTTTACCCCATCGCTCACATTTTCATACGAGTTA 832
QY 661 AACGCTTTAGGCTCTCGGGGAGCGGAGAAATGCGTTGTCATCATGATACAAATACCTATG 720
Db 833 AACGCTTTAGGCTCTCGGGGA-CGAGAAGAATGCGTTGTCATCATGATACAAATACCTATG 891
QY 721 AGCATCGTGGAGTCGGGAGGATGATGCTTCAACATCCATAGCTCTTTTCCAGTGG 780
Db 892 AGCATCGTGGAGTCGGGAGGATGATGCTTCAACATCCATAGCTCTTTTCCAGTGG 951
QY 781 TCATCACCAAGTCTTTAGGCTCACTCAAAACCATTCGCGAGTAATAGTGCATCATGTAGC 840
Db 952 TCATCACCAAGTCTTTAGGCTCACTCAAAACCATTCGCG-GTAAATAGTCATCATGTAGC 1010
QY 841 GACTCAGGCAACAACTGG 859
Db 1011 GACTCAGGCAAAACTGG 1029

RESULT 6
BX837992
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BX837992 795 bp mRNA linear EST 11-FEB-2004
BX837992 Arabidopsis thaliana cdna clone GSLTPGH21ZG06 5PRIM, mRNA sequence.
Arabidopsis thaliana
BX837992.1 GI:42532075
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 795)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished (2004)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen)
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1. .795
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

/ecotype="Col-0"
/db_xref="taxon:3702"
/clones="GSLTPGH21ZG06"
/issue_type="Hormone Treated Callus"
/clone_lib="Arabidopsis thaliana Hormone Treated Callus
Col-0"
ORIGIN
Query Match 52.5%; Score 598.2; DB 5; Length 795;
Best Local Similarity 95.6%; Pred. No. 4e-177;
Matches 626; Conservative 0; Mismatches 28; Indels 1; Gaps 1;
QY 1 ATGAGCTTCCGAAGTTTACTCCAGGAATGCGGTCTAGGCCACACCGTGTAGTTCACGGC 60
Db 141 ATGAGCTTCCGAAGTTTACTCCAGGAATGCGGTCTAGGCCACACCGTGTAGTTCACGGC 200
QY 61 GCCGCTCAACCGCTTAATAGTTTCAGACCCCTTTCAGCTGGTGGAGCTCCCGGAGAGCTG 120
Db 201 GCCGCTCAACCGCTTAATAGTTTCAGACCCCTTTCAGCTGGTGGAGCTCCCGGAGAGCTG 260
QY 121 CTTAGAGAAATCCTGATTTAGGTTGAGACTGTTGACGGGGCGGATTTGGCGTCGGCGGA 180
Db 261 CTTAGAGAAATCCTGATTTAGGTTGAGACTGTTGACGGGGCGGATTTGGCGTCGGCGGA 320
QY 181 AACGTGGTGGCTTGTGCGCGGCTTTGTGCTAGCTGGAGATTCTCAACGAGAGATTGTA 240
Db 321 AACGTGGTGGCTTGTGCGCGGCTTTGTGCTAGCTGGAGATTCTCAACGAGAGATTGTA 380
QY 241 GCTGTTCTCTGAATTCCTCTTAAATTTGACTTTTCCCTATCTCCCTCAAGCAGTCTGGTCCA 300
Db 381 GCTGTTCTCTGAATTCCTCTTAAATTTGACTTTTCCCTATCTCCCTCAAGCAGTCTGGTCCA 440
QY 301 AGAGATTTCTAGTTCAATGCTTTATAAAGCTAATCGAATACTCAATCGTATCATCTC 360
Db 441 AGAGATTTCTAGTTCAATGCTTTATAAAGCTAATCGAATACTCAATCGTATCATCTC 500
QY 361 TATCTCGGATTAACCTCTTTTGGCGGATACGAGGAGTTTCTTCTCTGCTTCTTAAG 420
Db 501 TATCTCGGATTAACCTCTTTTGGCGGATACGAGGAGTTTCTTCTCTGCTTCTTAAG 560
QY 421 CTGAAGCGCGCAACTTGCACTGATTACATCATCTCTTTGCGTTCAGACGATATCTCAAG 480
Db 561 CTGAAGCGCGCAACTTGCACTGATTACATCATCTCTTTGCGTTCAGACGATATCTCAAG 620
QY 481 AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTCTCTTGGAAACAAAATTCACGGTC 540
Db 621 AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTCTCTTGGAAACAAAATTCACGGTC 680
QY 541 TTTGATGTTAGTTCAGACCGGAGCGGAGATGCGAGAGCGGCTCTCTTAATTTTCATC 600
Db 681 TTTTATTTTATTCATACCTTATCATATATTTATATATTTATATATCTCTCTTCTTAATTTCAAC 740
QY 601 AAAGTTTCACTAGAGTTTCTCTCAGGGAA-GTTTACCCCATCGCTTCATTTTCATAC 654
Db 741 AAATCTTCACTATATATGCTCTCAATTAATGTTTACCCCATCTCTCTCAAAATTCATAC 795
RESULT 7
BX523610
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
BX523610 597 bp mRNA linear EST 19-MAR-2001
M39C9STM Arabidopsis developing seed Arabidopsis thaliana cdna
Clone M39C9 5', mRNA sequence.
BX523610
BX523610.1 GI:9781588
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 597)
White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.

161 GGGATTGGCCGTCGGCGGAAACGTGGTGGCTTGTGCGCGCTTGTGCTAGCTGGAGGA 220
Db |||||
191 GCGGATGGCCGTCACAGCCAGCGTGGTGGCTTGTGCGCGCTTGTGCTGGCGGC 250
Qy |||||
221 TTCTCACCAGGAGATTGAGTGTTCCTGAAATCTCTCTAAATGACTTTCCCTATCT 280
Db |||||
251 TACTTATGAAGAAACCGTGGTGTCCCTTGAGATCTCTTCTAAGTTGACTTTCCCACTCT 310
Qy |||||
281 CCTCAGCAGTCTGGTCCAGAGATTCTCTAGTTCATGCTTATTAAGCTTATAAGCTATCGAA 340
Db |||||
311 CTCTCAAGCAGCTGGTCCAGGGATTCTCTGGTTCATGCTTATCAACGTAATCGAA 370
Qy |||||
341 ATACTCAATCGTATCATCTCTCTCGAATTAACCTCTTTGACGGATAACGGAGT 400
Db |||||
371 TTACGGATCATATCATCTCTATCTCGGATTACCACTCTTTAACGGATGATGGAAGT 430
Qy |||||
401 TTCTCTCTGCTCTCTAAGCTGAAGCGCGCAACTTGCATGATTAATCATCTCTTTGTC 460
Db |||||
431 TTTTGTCTGCTGGTGAAGTTGAAGCACACAACCTTGTACGGATTACATTAATCTCTTTAC 490
Qy |||||
461 GTTCAGCAGTATCTCAAGAGAGCAACCGGTATCTTGGGAGATGAGATCGAACTTCC 520
Db |||||
491 GTTCTGATGATGTGCGAAGAGCAAGCTTATGTTGGCAAGTGAGATCGAACTTCC 550
Qy |||||
521 TTGGACAAATTCACGGTCTTTGATGTAGTCAG-----ACCGAGCAGCGAAGA 571
Db |||||
551 TAGGAACGAATTCACGTCTTTTGTGGAATCTGCTGCTTCAACGGGAGCGGAAAGT 610
Qy |||||
572 TGCAAGAAGCGCGCTCTCTTAATTTTCATCAAAAGTTTTCACCTAGAGTTCCTCAGGGAAGTT 631
Db |||||
611 TGAGAAGAAGCGGATCTTATATCCCGCAAAAGTTTCAGCAAAAGTTCTCTTTGGAAGTT 670
Qy |||||
632 ACCCCATCGTCAATTTTCATAGAGTTAAAGTCTTTAGGCTCTCGGAGCCAGGAGAA 691
Db |||||
671 ATCCTGCTGCTCATATAACATATGAGCTGAATGTCTTTAGGATCCCGGGACCAAGAAGA 730
Qy |||||
692 TGGTTGTCATCTAGGATCAATACCTATCAGCATCTGAGTCTGCGAGGAGTAGTACTT 751
Db |||||
731 TGCATATGCTTTATGGACACAATACCTTACAGCAATGAGGCTCAAGAGTAGCTTCAG 790
Qy |||||
752 CAACATCCATAAGCTCTTTTTCAGTCTGCTATCACCAGTCTTTAGTCTCTCACTCAAAAC 811
Db |||||
791 AACCATCAGAGTTTCCCTTACTCGGTACTCGGTCAACCTTATCCAGTCTCAGTCAAAAC 850
Qy |||||
812 CATTCGCGAGTATAGTCATCATGATAGCACTCAGGCAACACCTGGGAGATCCACAT 871
Db |||||
851 CATTCGCGAGTA-----GCTCAAGCCACCTGAAAGAAACACCAT 889
Qy |||||
872 TGGTCTGAGCAACAAGCTCCACGCTGGCATGAGCAGTTTACGTTCTGCTGCTTAAAT 931
Db |||||
890 TAGTCTGAGCAACAAGCACACGCTGGCAACGAGCTACGCTGCTGCTGCTGCTGAAAT 949
Qy |||||
932 TCCATGCTCGAGTCACAGTGGCTTCGGTTAAGAACTTTTCAGCTTTGTCGAG----- 982
Db |||||
950 TCCATGCGGTGTCACAGTAGCTCAGTGAAGAACTTTTCAGTCTGTCGCGCAGGAGCTA 1009
Qy |||||
983 TTAGTGACTGTGAAGCAGGCGACATCTGAGA-----GGATCATCTCCAGT 1030
Db |||||
1010 GCTGTGCGAGTGGCAGCGGAAATGTCACCGGAGAGCGAGCGGAGTATATATTCAGT 1069
Qy |||||
1031 TTGGGAAGTTGGGAAGGACATCTTTTACCATGATTTATGATATCGATTTCTGCGTTTC 1090
Db |||||
1070 TTGGGAAGTGGCGGAAGAATGTTTACGATGATTTATGATATACCGATCTCAGCTTTC 1129
Qy |||||
1091 AAGCTTTTGTATCTGCTCAGCAGTTTTGAACACCAAGATTCGCTGTGAA 1140
Db |||||
1130 AGCTTTTGCAATTGCTTCAGCGCTCTGAGACTAGAACTCGCTTGTGAA 1179

RESULT 9
AV527654 516 bp mRNA linear EST 19-FEB-2004
LOCUS
DEFINITION AV527654 Arabidopsis thaliana aboveground organs two to six-week

old Arabidopsis thaliana cDNA clone AP242a03R 5', mRNA sequence.
AV527654 GI:8687182
AV527654.1
EST.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 516)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7 (3), 175-180 (2000)
10907847
Contact: Brika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp. URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..516
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="AP242a03R"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana aboveground organs two to
six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
ORIGIN
Query Match 45.3%; Score 516; DB 1; Length 516;
Best Local Similarity 100.0%; Pred. No. 3.4e-151;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 40 CCACACCGTGTAGTTTCAGCGCGCGCCCTCAACCGCTAATAGTTTCAGACCCCTTCAGCTGG 99
Db 1 CCACACCGTGTAGTTTCAGCGCGCGCCCTCAACCGCTAATAGTTTCAGACCCCTTCAGCTGG 60
Qy 100 TCGGAGCTCCCGAGGAGCTGCTTAGAGAAATCTCTGATTAGGGTTGAGACTGTTGACGCG 159
Db 61 TCGGAGCTCCCGAGGAGCTGCTTAGAGAAATCTCTGATTAGGGTTGAGACTGTTGACGCG 120
Qy 160 GCGGATTGGCCGCTCGCGCGAAACGTTGGTGGCTTTGCGCGGCTTTGTCGTAGCTGGAGG 219
Db 121 GCGGATTGGCCGCTCGCGCGAAACGTTGGTGGCTTTGCGCGGCTTTGTCGTAGCTGGAGG 180
Qy 220 ATTCTCACCAAGGAGATTGTAGCTGTTCTGAATTCCTCTAAATTTGACTTCCCTATC 279
Db 181 ATTCTCACCAAGGAGATTGTAGCTGTTCTGAATTCCTCTAAATTTGACTTCCCTATC 240
Qy 280 TCCCTCAAGCAGCTGGTCCAAAGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAATCGA 339
Db 241 TCCCTCAAGCAGCTGGTCCAAAGAGATTCTCTAGTTCAATGCTTTTATAAAACGTAATCGA 300
Qy 340 AATACTCAATCTATCATCTCTATCTCGGATTAACCTACCTTTTGACGGAATACCGGAAG 399
Db 301 AATACTCAATCTATCATCTCTATCTCGGATTAACCTACCTTTTGACGGAATACCGGAAG 360
Qy 400 TTTCTTTTGTGCTTCTAAGCTGAAGCGGCAACTTGACACTGATTACATCATCTCTTTG 459
Db 361 TTTCTTTTGTGCTTCTAAGCTGAAGCGGCAACTTGACACTGATTACATCATCTCTTTG 420
Qy 460 CGTTTCAGACGATATCTCTCAAGAGAGCAACGGTATCTTTGGGAGAAATGAGATCGAACTTC 519
Db 421 CGTTTCAGACGATATCTCTCAAGAGAGCAACGGTATCTTTGGGAGAAATGAGATCGAACTTC 480
Qy 520 CTTGGAAACAAAATTACCGGTCTTTTGATGTTAGTCAG 555

QY	361	TATCTCGGATTAACTACCTCTTTGACGGATAACGGGAAGTTTCTTCTTCTGCTGCTTCTTAAG	420
Db	532	TATCTCGGATTAACTACCTCTTTGACGGATAACGGGAAGTTTCTTCTTCTGCTGCTTCTTAAG	591
QY	421	CTGAAGCGCGCAACTTGCACATGATTACATCATCTCTTTGGCGTTCCAG	466
Db	592	CTGAAGCGCGCAACTTGCACATGATTACATCATCTCTTTGGCGTTCCAG	637
RESULT 11			
LOCUS	AU2371181		
DEFINITION	AU2371181 RAFL16 Arabidopsis thaliana cDNA clone RAFL16-03-H24 5', mRNA sequence.		
ACCESSION	AU2371181		
VERSION	AU2371181.1	GI:19876350	
KEYWORDS	EST.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 598)		
AUTHORS	Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.		
TITLE	Large scale analysis of Arabidopsis full-length cDNA		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-3060 Email: msaki@rct.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.		
FEATURES			
source	Location/Qualifiers		
	1..598		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/clone="RAFL16-03-H24"		
	/lab_host="DH10B"		
	/clone_lib="RAFL16"		
	/note="Site_1: BamHI; Site_2: SalI; dark-grown"		
ORIGIN			
Query Match	39.4%;	Score 449;	DB 1; Length 598;
Best Local Similarity	99.3%;	Pred. No. 5.7e-130;	
Matches	449;	Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	1	ATGAGCTTCCGAGTTTACTCCAGGAATGCGGTCTAGGCCACACCGTGTAGTTTCAGCC	60
Db	147	ATGAGCTTCCGAGTTTACTCCAGGAATGCGGTCTAGGCCACACCGTGTAGTTTCAGCC	206
QY	61	GC CGCCTCAACCGCTTAATAGTTTCAGACCTTTTCAGCTGGTCCGAGCTCCCGAGGAGCTG	120
Db	207	GC CGCCTCAACCGCTTAATAGTTTCAGACCTTTTCAGCTGGTCCGAGGAGCTG	266
QY	121	CTTAGAGAAATCCTGATTAGGTTGAGACTGTTCACGCGCGCGGATTTGGCGTCCGCGCGA	180
Db	267	CTTAGAGAAATCCTGATTAGGTTGAGACTGTTCACGCGCGCGGATTTGGCGTCCGCGCGA	326
QY	181	AACGTGGTGGCTTGTGCGCGGTTTGTGCTAGCTGGAGGATTTCTCAACAAGGAGATTGTA	240
Db	327	AACGTGGTGGCTTGTGCGCGGTTTGTGCTAGCTGGAGGATTTCTCAACAAGGAGATTGTA	386

```
QY 241 GCTGTTCTCGAATTCCTCTAAATGACTTTCCTATCTCCCTCAAGCAGTCTGGTCCA 300
Db 387 GCTGTTCTCGAATTCCTCTAAATGACTTTCCTATCTCCCTCAAGCAGTCTGGTCCA 446
QY 301 AGAGATTCTCTAGTTCATGCTTTTATAAAACGTAATCGAAATACCTCAATCGTATCATCTC 360
Db 447 AGAGATTCTCTAGTTCATGCTTTTATAAAACGTAATCGAAATACCTCAATCGTATCATCTC 506
QY 361 TATCTCGGATTAACTACCTCTTTGACGGATAACGGGAAGTTTCTTCTTGCTGCTTCTTAAG 420
Db 507 TATCTCGGATTAACTACCTCTTTGACGGATAACGGGAAGTTTCTTCTTGCTGCTTCTTAAG 566
QY 421 CTGAAGCGCGCAACTTGCACTGATTACATCAT 452
Db 567 CTGAAGCGCGCAACTTGCACTGATTACATCAT 598

RESULT 12
BU635249
LOCUS
DEFINITION
003805 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA
sequence.
ACCESSION
BU635249
VERSION
BU635249.1 GI:23302504
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 704)
Lundegaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S.
and Welinder,K.G.
EST sequencing of Erysiphe cichoracearum infected Arabidopsis
Plants
Unpublished (2002)
Contact: Karen G. Welinder
Institut for Biotechnologi
Aalborg Universitet
Sohngaardsbolnavej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.au.dk.
Location/Qualifiers
1. .704
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/dev_stage="Plant 3 weeks old, three days post infection"
/clone_lib="Infected Arabidopsis Leaf"
/notes="Organ: Leaf; Vector: pBluescript; Mixed cDNA
library of Arabidopsis and E. cichoracearum infected leaf
from three weeks old Arabidopsis plants. Plants were
harvested 3 days after infection and mRNA oligo dt
selected."

FEATURES
source
Query Match 37.1%; Score 422.6; DB 5; Length 704;
Best Local Similarity 76.8%; Pred. No. 1.3e-121;
Matches 534; Conservative 0; Mismatches 154; Indels 9; Gaps 1;

ORIGIN
QY 58 GCCGCCGCTCAACCGCTAATAGTTCAGACCCCTTCAGCTGTCGGAGCTCCGCGAGGAG 117
Db 7 GCTGCCGATTCCTCTGCTGTCATCGCAAGATTATCGCTGGTCAGAGATTCTCTGAAGAG 66
QY 118 CTGCTTAGAGAAATCTCTGATTAGGTTGAGACTGTTGACCGCGCGCAATTGGCCGCTCGCG 177
Db 67 CTTCTTAGGAGAGATTCTGATTCTGTTGTTGAACGCGCGACGTCGCGGATGGCCGTCACGA 126
QY 178 CGAAACGTGCTGCTGCGCGCTTTGTCGTAGCTGGAGGATTCTCACCAAGGAGATT 237
```

```
Db 127 CCAGCGTGTGTGGCTTGTGCGGGGTTTTGTGCTGGCTGGCGGCTACTTATGAACGAACC 186
QY 238 GTAGCTGTTCTGAATTCCTCTCTAAATTTGACTTTTCCCTATCTCCCTCAAGCAGTCTGGT 297
Db 187 GTGCTGTGCTCGAGATCTCTTTAAGTTGACTTTTCCCATCTCTCTCAAGCAGCTGGT 246
QY 298 CCAAGAGATTCTCTAGTTCATGCTTTATAAAACGTAATCGAAATACCTCAATCGTATCAT 357
Db 247 CCAAGGATTCTCTAGTTCATGCTTTATAAAACGTAATCGAAATACCTCAATCATCAT 306
QY 358 CTCTATCTCGGATTAACTACCTCTTTGACGGATAACGGGAAGTTTCTTCTTGCTGCTTCT 417
Db 307 CTCTATCTCGGATTAACTACCTCTTTGACGGATAACGGGAAGTTTCTTCTTGCTGCTGCT 366
QY 418 AAGCTGAAGCGCGCAACTTGCACTGATTACATCATCTCTTTGCGTTTCAGACGATATCTCA 477
Db 367 AAGTTGAAGCACACAACTTGACGGATTACATCATCTCTTTACGTTCTGATGATATGCG 426
QY 478 AAGAGAAGCAACGGCTATCTTTGGGAGAAATGAGATCGAACTTCTCTTGGAAACAAATTCACG 537
Db 427 AGAAGAACCAAGCTTATGTTGGCAAGTGAGATCGAACTTCTCTAGGAACGAAATTCCT 486
QY 538 GTCTTTGATGTGTAGTCTAG-----ACCGAGCAGCGAAGATCGAAGAGCGCTCT 588
Db 487 GTCTTTGATGTGTAGTCTAGTCTTCAACGGGAGCGCAAAAGTTTGAGAAAGAGCGGATCT 546
QY 589 TCTAATTTTCATCAAGTTTTCACCTAGAGTTCTCTCAGGGAAGTTTACCCATCTCACATT 648
Db 547 TATAATCCCGCAAAAGTTTCAGCAAAAGTTCTCTTTGGAAGTTATCTCTGCTCTCATATC 606
QY 649 TCATACGAGTTTAAACGCTCTTAGGCTCTCGGGACCGAGAGAAGTTCGTTGCAATCATGGAT 708
Db 607 ACATATGAGTCAATGTTCTTAGGATCCCGGGGACCACCAAGAGATGCAATGCTTATGGAC 666
QY 709 ACAATPACTATGAGCATCTGTGGAGTCGCGAGGAGTAG 745
Db 667 ACAATPACTATGAGCATCTGTGGAGTCGCGAGGAGTAG 703

RESULT 13
F13831
LOCUS
DEFINITION
ATT54753 Versaillies-VB Arabidopsis thaliana cDNA clone VBVOC09,
similar to phosphodiesterase, mRNA sequence.
ACCESSION
F13831
VERSION
F13831.1 GI:757555
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 420)
CNRS.
The Arabidopsis thaliana transcribed genome: the GDR cDNA program
Unpublished (1996)
Contact: Desprez T., Amselem J., Chiapello H., Rouze P., Caboche
M., Hoite H.
INRA Versailles
Laboratoire de Biologie Cellulaire
Route de Saint-Cyr, 78026 Versailles Cedex, France
Email: thierry@versaillies.inra.fr.
Location/Qualifiers
1. .420
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/scot_type="Columbia"
/db_xref="taxon:3702"
/clone="VBVOC09"
/tissue_type="whole seedlings"
/dev_stage="in vitro-grown etiolated seedlings, 5 days old"
/clone_lib="Versailles-VB"
/note="Vector: pBluescript"

FEATURES
source
Query Match 37.1%; Score 422.6; DB 5; Length 704;
Best Local Similarity 76.8%; Pred. No. 1.3e-121;
Matches 534; Conservative 0; Mismatches 154; Indels 9; Gaps 1;
```


ORIGIN	selected."	ORIGIN	
Query Match	34.0%; Score 388; DB 8; Length 420;	Query Match	29.2%; Score 333; DB 5; Length 728;
Best Local Similarity	98.3%; Pred. No. 9.8e-111;	Best Local Similarity	74.3%; Pred. No. 3e-93;
Matches 413; Conservative	0; Mismatches 5; Indels 2; Gaps 2;	Matches 472; Conservative	0; Mismatches 135; Indels 28; Gaps 3;
QY	262 AAATTGACATTTCCCTATCTCCCTCAAGCAGTCTGGTCCAAAGAGATTCTCTAGTTCAATGC 321	QY	1 ATGACGTTCCGAGATTACTCCAGGAAATGCGGTCTAGGCCACACCGGTGTAGTTCACGCC 60
Db	1 AAATTGACATTTCCCTATCTCCCTCAAGCAGTCTGGTCCAAAGAGATTCTCTAGTTCAATGC 60	Db	77 ATGACCTTTACGTAGCTTAATCTCTTGAGATGCGTTCGAGACCGCATCTGTGTGTCACGAC 136
QY	322 TTATATAACGTAATCGAATACTCAATCGTATCATCTCTATCTCGGATTAACCTACTCT 381	QY	61 -----GCCGCTCAACCGCTCAATAGTTTCAGACCCCTTTTCAGCTGTGTCG 102
Db	61 TTATATAACGTAATCGAATACTCAATCGTATCATCTCTATCTCGGATTAACCTACTCT 120	Db	137 CTTGCGCGCGCGCAGCTGCCGATTCCACTTCTGTGTCTATCGCAAGATTATCGCTGTGTCA 196
QY	382 TTGACGGATAACGGGAAGTTTCTTCTTGTCTCTTCAAGCTGAAGCGCGCAACTTGCACT 441	QY	103 GAGCTCCCGAGAGAGCTGCTTAGAGAAATCTCTGATTAGGTTTGAAGACTGTTTGAACGGCGC 162
Db	121 TTGACGGATAACGGGAAGTTTCTTCTTGTCTCTTCAAGCTGAAGCGCGCAACTTGCACT 180	Db	197 GAGATTCTCTGAAGAGCTTCTTAGGGAGATTCTGATTCTGTTCGTTGAAGCGCGGAGCGGTGGC 256
QY	442 GATTACATCATCTCTTGGCTTCAGACGATATCTCAAGAGAGACGCGTATCTTGGG 501	QY	163 GATTGCGCGCTCGCGG-GAAACGTGGTCTGTGCGCGCGTGTTCGTAGCTGGAGGAT 221
Db	181 GATTACATCATCTCTTGGCTTCAGACGATATCTCAAGAGAGACGCGTATCTTGGG 240	Db	257 GGATGCGCGCTCAGACGCGCAGCGTGTGCTGTGCGCGCGTGTGTCTGCTGGCTGGCGGCT 316
QY	502 AGAATGAGATCGAATCTCTTGGAAACAAATTCACGCTCTTTGATGGTAGTCAGACCGGA 561	QY	222 TCTCAACAAAGAGATTGTAGCTGTTCTCTGAATTTCTCTCTAAATTTGACTTTCCCTATCTC 281
Db	241 AGAATGAGATCGAATCTCTTGGAAACAAATTCACGCTCTTTGATGGTAGTCAGACCGGA 300	Db	317 ACTTATGAACGAACCGTGTCTCTGAGATCTCTTCTAAGTTGACTTTTCCCATCTC 376
QY	562 GCAGCGAAGATGAG-AAGAGCGCTCTTCTTAATTTTCATCAAAAGTTTCACCTAGAGTTCC 620	QY	282 CTTCAAGCAGTCTGGTCCAAAGAGATTCTCTAGTTTCAATGCTTTTATAAAGTAAATCGAAA 341
Db	301 GCAGCGAAGATGAGAGAGCGCTCTTCTTAATTTTCATCAAAAGTTTCACCTAGAGTTCC 360	Db	377 TCTCAAGCAGCTGGTCCAGGGATTCACTGTTTCAATGCTTTTATCAAAAGTAAATCGAAT 436
QY	621 TCAGGGAAGTTACCCCATCGC-TCACATTTTCATACGAGTTTAAACGTTCTTAGGCTCTCGGG 679	QY	342 TACTCAATCGTATCATCTCTATCTCGGATTACTACTCTTTTACCGGATAACGGGAGATT 401
Db	361 TCAGGGAAGTTACCCCATCGCTTCACATTTTCATACGAGTTTAAACGTTCTTAGGCTCTCGGG 420	Db	437 TAGGCAATCATATCATCTCTATCTCGGATTAAACCAACTCTTTTAAACGAGATGAGGAGATT 496
RESULT 14		QY	402 TCTTTCTGCTGCTTCTTAAGCTGAAGCGCGCAACTTGCACCTTGATTCATCATCTCTTTGCG 461
LOCUS	BU635253	Db	497 TTTGCTTGTGCTGTAGTTGGAAGCACAACTTGTACGGATTACATATCTCTTTTACG 556
DEFINITION	018803 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA	QY	462 TTCAGACGATATCTCAAGAGAGAACCGGTAATCTTTGGGAGAAATGAGATCGAATTCCT 521
ACCESSION	BU635253	Db	557 TTCTGATGATATGTCGAGAAGAAGCAAGCTATGTTGGCAAGTGGATCGAATTCCT 616
VERSION	BU635253.1 GI:23302508	QY	522 TGGAAACAAAATTCACGCTCTTTGATGGTAGTCAG-----ACCGGAGCAGCGAAGAT 572
KEYWORDS	EST.	Db	617 AGGAACGAAATTCACCTGTCTTTGATGGAATCTGTCCCTTCAACGGGAGCCGCAAGTT 676
SOURCE	Arabidopsis thaliana (thale cress)	QY	573 GCAGAAAGACGCGCTCTTCTTAATTTTCATCAAAAGTTT 607
ORGANISM	Arabidopsis thaliana	Db	677 GAGAAAGACCGCATCTTATATATCCCGCAAAAGTTT 711
REFERENCE	Lundgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S. and Welinder,K.G.	RESULT 15	
AUTHORS	EST sequencing of Erysiphe cichoracearum infected Arabidopsis plants	T44322	
TITLE	Unpublished (2002)	LOCUS	T44322
JOURNAL	Contact: Karen G. Welinder	DEFINITION	7585 Lambda-PRL2 Arabidopsis thaliana cDNA clone 124E22T7, mRNA
COMMENT	Institut for bioteknologi Aalborg Universitet Sohngaardsholmsvej 49, 9000 Aalborg, Denmark Tel: +45 96358467 Fax: +45 98141808 Email: kgw@bio.au.dk.	ACCESSION	T44322
FEATURES	Location/Qualifiers	VERSION	T44322.1 GI:2759115
source	1. .728	KEYWORDS	EST.
	/organism="Arabidopsis thaliana"	SOURCE	Arabidopsis thaliana (thale cress)
	/mol_type="mRNA"	ORGANISM	Arabidopsis thaliana
	/scot_type="Columbia"	REFERENCE	1 (bases 1 to 358)
	/db_xref="taxon:3702"	AUTHORS	Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel,B. and Somerville,C.
	/dev_stage="Plant 3 weeks old, three days post infection"	TITLE	Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
	/clone_lib="Infected Arabidopsis Leaf"	JOURNAL	Plant Physiol. 106, 1241-1255 (1994)
	/notes="Organ: Leaf; Vector: pBluescript; Mixed cDNA library of Arabidopsis and E. cichoracearum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dr	PUBMED	7846151

COMMENT

On Jan 7, 1998 this sequence version replaced gi:948641.

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ibm.cl.msu.edu
Seq primer: T7.

FEATURES

source

Location/Qualifiers
1..358
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="124E227"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA. "

ORIGIN

Query Match 29.1%; Score 331.2; DB 8; Length 358;
Best Local Similarity 93.9%; Pred. No. 8.6e-93;
Matches 336; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 149 CTGTTGACGGCGCGGATTGGCCCTCGCGGCGAAACGTGGTGGCTTGTGCGCGGCTTTGTC 208
DB 1 CTGTTGACGGCGCGGATTGGCCCTCGCGGCGAAACGTGGTGGCTTGTGCGCGGCTTTGTC 60
QY 209 GTAGCTGGAGGATTCACCAAGGAGATTGTAGCTGTTCTGTAATTCCTCTAAATTGA 268
DB 61 GTAGCTGGAGGATTCCTNACCAGGAGATTGTAGCTGTTCTGTAATTCCTCTAAATTGA 120
QY 269 CTTTCCTATCTCCCTCAAGCAGTCTGGTCCAGAGATTCTCTAGTTCAATGCTTTATAA 328
DB 121 CTTTCCTATCTCCCTCAAGCAGTCTGGTCCAGAGATTCTCTAGTTCAATGCTTTATAA 180
QY 329 AACGTAATCGAATATCTCAATCGTATCATCTCTATCTCGGATTAACCTCTTTGACGG 388
DB 181 AACGTAATCGAATATCTCAATCGTATCATCTCTATCTCGGATTAACCTCTTTGACGG 240
QY 389 ATAACGGGAAGTTTCTTTCTGCTTTCTTAAGCTGAAGCGCGCAACTTGCACTGATTACA 448
DB 241 ATAACGGGAAGTTTCTTTCTGCTTTCTTAAGCTGAAGCGCGCAACTTGCACTGATTACA 300
QY 449 TCATCTCTTTGCGTTTCAGAGATATCTCAAGAGAGACGCGTATCTTGGGAGAT 506
DB 301 TCATCTCTTTGCGTTTCAGAGATATCTCAAGAGAGACGCGTATCTTGGGAGAT 358

Search completed: December 26, 2005, 06:49:27
Job time : 5224 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 03:25:36 ; Search time 267 Seconds
(without alignments)
7589.589 Million cell updates/sec

Title: US-10-763-042-20
Perfect score: 1140
Sequence: 1 atgagctccgaagtact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/H COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/PCUS COMB.seq.*
- 7: /cgn2_6/prodata/1/ina/PP COMB.seq.*
- 8: /cgn2_6/prodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	6.3	299	3	US-09-313-294A-6623
2	66.8	5.9	1426	2	US-08-630-592-6
3	66.8	5.9	1426	2	US-08-714-991-6
4	66.8	5.9	1426	3	US-09-032-365A-7
5	66.8	5.9	1518	3	US-08-955-918C-11
6	66.8	5.9	1518	3	US-08-697-766A-11
7	66.8	5.9	1890	3	US-09-032-365A-61
8	66.8	5.9	1936	3	US-09-032-365A-59
9	66.8	5.9	2040	2	US-08-631-200-7
10	66.8	5.9	2040	2	US-08-929-553-7
11	66.8	5.9	2040	2	US-08-936-707A-7
12	66.8	5.9	2040	2	US-08-936-706A-1
13	66.8	5.9	2040	2	US-09-248-203-7
14	66.8	5.9	2040	3	US-09-406-071-7
15	66.8	5.9	2040	3	US-08-955-918C-9
16	66.8	5.9	2040	3	US-08-697-766A-9
17	66.8	5.9	2040	3	US-09-814-986-7
18	66.8	5.9	2088	3	US-09-032-365A-64
19	66.8	5.9	2109	3	US-09-032-365A-63
20	66.8	5.9	2112	3	US-09-032-365A-56
21	66.8	5.9	2119	2	US-08-630-592-1
22	66.8	5.9	2119	2	US-08-714-991-1
23	66.8	5.9	2119	3	US-09-032-365A-1
24	66.8	5.9	2119	3	US-09-032-365A-1

25	66.8	5.9	2368	3	US-09-032-365A-57	Sequence 57, Appli
26	66.8	5.9	2434	2	US-08-630-592-3	Sequence 3, Appli
27	66.8	5.9	2434	2	US-08-714-991-3	Sequence 3, Appli
28	66.8	5.9	2434	3	US-09-032-365A-3	Sequence 26, Appli
29	66.8	5.9	3060	2	US-08-714-991-26	Sequence 9, Appli
30	66.8	5.9	3268	3	US-09-032-365A-9	Sequence 8, Appli
31	65.2	5.7	1512	3	US-08-955-918C-8	Sequence 8, Appli
32	65.2	5.7	1512	3	US-08-697-766A-8	Sequence 8, Appli
33	65.2	5.7	1801	3	US-08-955-918C-6	Sequence 6, Appli
34	65.2	5.7	1801	3	US-08-697-766A-6	Sequence 6, Appli
35	65.2	5.7	1804	2	US-08-631-200-1	Sequence 1, Appli
36	65.2	5.7	1804	2	US-08-829-553-1	Sequence 1, Appli
37	65.2	5.7	1804	2	US-08-922-267A-1	Sequence 1, Appli
38	65.2	5.7	1804	2	US-08-936-707A-1	Sequence 1, Appli
39	65.2	5.7	1804	2	US-08-936-706A-1	Sequence 1, Appli
40	65.2	5.7	1804	3	US-09-248-203-1	Sequence 1, Appli
41	65.2	5.7	1804	3	US-09-406-071-1	Sequence 1, Appli
42	65.2	5.7	1804	3	US-09-814-986-1	Sequence 1, Appli
43	65.2	5.7	2570	3	US-09-270-767-13498	Sequence 13498, A
44	64.6	5.7	1482	3	US-09-032-365A-16	Sequence 16, Appli
45	62.2	5.5	1743	3	US-09-032-365A-18	Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-09-313-294A-6623
; Sequence 6623, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ico, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6623
; LENGTH: 299
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700352120H1
; NAME/KEY: unsure
; LOCATION: 102, 134, 269
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6623

Query Match 6.3%; Score 72; DB 3; Length 299;
Best Local Similarity 72.1%; Pred. No. 1.3e-13;
Matches 93; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1012 GAGAGATCATCTCCAGTTTCGGAAGTTGGGAAGACATGTTTACCATGATTATGA 1071
Db 2 GAAAGGTGATCCTGCAATTTGGGAAGTGGTGAAGACATGTTTACCATGACTACCGT 61
QY 1072 TATCCGATTTTCGGTTTCAAGGTTTCTATCTGCTGAGCAGTTTTCGAACCAAGAA 1131
Db 62 TACCACCTGTCAGCCTTCAGGCTTCGCGATCTGCGTTCGAGCTTCGACCAAGCTG 121
QY 1132 GCCTGTGAA 1140
Db 122 GCCTGCGAA 130

RESULT 2
US-08-630-592-6
; Sequence 6, Application US/08630592
; Patent No. 5770432

GENERAL INFORMATION:
APPLICANT: Nishina, Patsy
APPLICANT: No. 5770432enTrauth, Konrad
APPLICANT: Naggert, Juergen
APPLICANT: No. 5770432th, Michael
TITLE OF INVENTION: Obesity Associated Genes
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 941114187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,592
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
TELEPHONE: (415) 7811989
TELEFAX: (415) 3983249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1426 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-630-592-6

Query Match 5.9%; Score 66.8; DB 2; Length 1426;
Best Local Similarity 57.1%; Pred. No. 1.9e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;
QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTACGTTGCTGGTCTTAAAT 931
Db 1157 TCGAGCTGCAAAACAAGACACCTGTCTGGAATGATGACACACAGTCTATGACTCAACT 1216
QY 932 TCCATGGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1217 TCCATGGGCGGTCACACAGGCTCCGTTGAAGAACTTCCAGATCATCCATGGCAATGACC 1276
QY 992 GTGAAGCAGGCGACACATCTGAGAGGATCATACTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
Db 1277 -----CGGACTACATCTGTGATGAGTTTGGCGGTTAGCAGAGGATG 1318
QY 1052 TGTATTACCATGATTATGAGATATCCGATTTCTCGTTTCAAGCGTTTGTCTATCTGCTGA 1111
Db 1319 TGTTCACCATGATTACAACTACCCGCTGTGTGCACTGTCAGGCTTTGGCAATTGCCCTGT 1378
QY 1112 GCAGTTTGAACACAGAAATTCGCTGTGA 1139
Db 1379 CCAGCTTCGACAGCAAGCTGGCGTGGCA 1406

RESULT 3
US-08-714-991-6
Sequence 6, Application US/08714991
Patent No. 5776762
GENERAL INFORMATION:
APPLICANT: NISHINA, Michael
APPLICANT: NISHINA, Patsy
APPLICANT: No. 5776762en-Trauth, Konrad

APPLICANT: NAGGERT, Juergen
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,991
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-59504-1/PJS
TELEPHONE: 415-494-8700
TELEFAX: 415-494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1426 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-714-991-6
Query Match 5.9%; Score 66.8; DB 2; Length 1426;
Best Local Similarity 57.1%; Pred. No. 1.9e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;
QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTACGTTGCTGGTCTTAAAT 931
Db 1157 TCGAGCTGCAAAACAAGACACCTGTCTGGAATGATGACACACAGTCTATGACTCAACT 1216
QY 932 TCCATGGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1217 TCCATGGGCGGTCACACAGGCTCCGTTGAAGAACTTCCAGATCATCCATGGCAATGACC 1276
QY 992 GTGAAGCAGGCGACACATCTGAGAGGATCATACTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
Db 1277 -----CGGACTACATCTGTGATGAGTTTGGCGGTTAGCAGAGGATG 1318
QY 1052 TGTATTACCATGATTATGAGATATCCGATTTCTCGTTTCAAGCGTTTGTCTATCTGCTGA 1111
Db 1319 TGTTCACCATGATTACAACTACCCGCTGTGTGCACTGTCAGGCTTTGGCAATTGCCCTGT 1378
QY 1112 GCAGTTTGAACACAGAAATTCGCTGTGA 1139
Db 1379 CCAGCTTCGACAGCAAGCTGGCGTGGCA 1406

RESULT 4
US-09-032-365A-7
Sequence 7, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:
APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: Naggert, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
TITLE OF INVENTION: NEUROSENSORY DEFECTS
NUMBER OF SEQUENCES: 67

```
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Bozicevic & Reed, LLP
/ STREET: 285 Hamilton Avenue, Suite 200
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94301
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/032,365A
/ FILING DATE:
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sherwood, Pamela J
/ REGISTRATION NUMBER: 36,677
/ REFERENCE/DOCKET NUMBER: SEQ-2CIP2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-327-3400
/ TELEFAX: 650 327-3231
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1426 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ US-09-032-365A-7

Query Match 5.9%; Score 66.8; DB 3; Length 1426;
Best Local Similarity 57.1%; Pred. No. 1.9e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTGTACGTTGCTGGTGTAAATT 931
Db 1157 TCAGCTGCAAAACAGACACCTGCTGGATGATGACACACAGTCTATGTACTCACT 1216
QY 932 TCATGCTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1217 TCATGGGCGGTGACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1276
QY 992 GTGAAGCAGGGCAGACATCTGAGAGGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1277 -----CGGACTACATCGTATGATGAGTTGGCGGTAGCAGAGGATG 1318
QY 1052 TGTTTACCATGGATTATGGATATCCGATTTCTCGGTTTCAAGGGTTTGTCTATCTGCTGA 1111
Db 1319 TGTTACCATGGATTACAACTCCGCTGTGTGCACTGCAAGGCTTTGCCATTTGCCCTGT 1378
QY 1112 GCAGTTTTGAAACCAAGATTGCTGTGA 1139
Db 1379 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1406

RESULT 5
US-08-955-918C-11
; Sequence 11, Application US/08955918C
; Patent No. 6268130
; GENERAL INFORMATION:
; APPLICANT: Klymn, Patrick, and Moore, Karen
; TITLE OF INVENTION: RP Compositions and Therapeutic and
; NUMBER OF INVENTIONS: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts

/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Bozicevic & Reed, LLP
/ STREET: 285 Hamilton Avenue, Suite 200
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94301
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/032,365A
/ FILING DATE:
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sherwood, Pamela J
/ REGISTRATION NUMBER: 36,677
/ REFERENCE/DOCKET NUMBER: SEQ-2CIP2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-327-3400
/ TELEFAX: 650 327-3231
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1426 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ US-09-032-365A-7

Query Match 5.9%; Score 66.8; DB 3; Length 1426;
Best Local Similarity 57.1%; Pred. No. 1.9e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTGTACGTTGCTGGTGTAAATT 931
Db 1157 TCAGCTGCAAAACAGACACCTGCTGGATGATGACACACAGTCTATGTACTCACT 1216
QY 932 TCATGCTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1217 TCATGGGCGGTGACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1276
QY 992 GTGAAGCAGGGCAGACATCTGAGAGGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1277 -----CGGACTACATCGTATGATGAGTTGGCGGTAGCAGAGGATG 1318
QY 1052 TGTTTACCATGGATTATGGATATCCGATTTCTCGGTTTCAAGGGTTTGTCTATCTGCTGA 1111
Db 1319 TGTTACCATGGATTACAACTCCGCTGTGTGCACTGCAAGGCTTTGCCATTTGCCCTGT 1378
QY 1112 GCAGTTTTGAAACCAAGATTGCTGTGA 1139
Db 1379 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1406

RESULT 5
US-08-955-918C-11
; Sequence 11, Application US/08955918C
; Patent No. 6268130
; GENERAL INFORMATION:
; APPLICANT: Klymn, Patrick, and Moore, Karen
; TITLE OF INVENTION: RP Compositions and Therapeutic and
; NUMBER OF INVENTIONS: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts

/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02109-1875
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/955,918C
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/697,766
/ FILING DATE: 29-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mandragouras, Amy E.
/ REGISTRATION NUMBER: 36,207
/ REFERENCE/DOCKET NUMBER: MNI-007CPDV2CPA
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617)227-7400
/ TELEFAX: (617)227-5941
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1518 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..1519
/ US-08-955-918C-11

Query Match 5.9%; Score 66.8; DB 3; Length 1518;
Best Local Similarity 57.1%; Pred. No. 2e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTGTACGTTGCTGGTGTAAATT 931
Db 1268 TCAGCTGCAAAACAGACACCTGCTGGATGATGACACACAGTCTATGTACTCACT 1327
QY 932 TCATGCTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1328 TCATGGGCGGTGACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1387
QY 992 GTGAAGCAGGGCAGACATCTGAGAGGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1388 -----CGGACTACATCGTATGATGAGTTGGCGGTAGCAGAGGATG 1429
QY 1052 TGTTTACCATGGATTATGGATATCCGATTTCTCGGTTTCAAGGGTTTGTCTATCTGCTGA 1111
Db 1430 TGTTACCATGGATTACAACTACCCGCTGTGTGCACTGCAAGGCTTTGCCATTTGCCCTGT 1489
QY 1112 GCAGTTTTGAAACCAAGATTGCTGTGA 1139
Db 1490 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1517

RESULT 6
US-08-697-766A-11
; Sequence 11, Application US/08697766A
; Patent No. 6399760
; GENERAL INFORMATION:
; APPLICANT: Gimeno, Carlos; Klymn, Patrick; and Moore, Karen J.
; TITLE OF INVENTION: RP Compositions and Therapeutic and
; NUMBER OF INVENTIONS: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
```

; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/697,766A
; FILING DATE: 29-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1518
; US-08-697-766A-11

Query Match 5.9%; Score 66.8; DB 3; Length 1518;
Best Local Similarity 57.1%; Pred. No. 2e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY	872	TGCTGCTGAGCAACAAGCTCCACGGTGGCATGACGAGTTACGTTCTGGTCTTAAAT	931
Db	1268	TCGAGCTGCAAAACAGACACCTGCTGGAATGACACACAGTCTATGACTCACT	1327
QY	932	TCCATGCTCGAGTCACAGTGGCTTCGGTTAAGAACTTCCAGCTTTGGCAGTTAGTGACT	991
Db	1328	TCCATGGCGCGTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGCGCAATGACC	1387
QY	992	GTGAAGCAGGCGACAGATCTGAGAGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA	1051
Db	1388	-----CGGACTACATCGTGATGCGAGTTTGGCGGCTAGCAGAGGATG	1429
QY	1052	TGTTTACCATGGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGCTATCTGCTGA	1111
Db	1430	TGTTTACCATGGATTATCACTCCCGCTGTGTGCACTGCAGGCTTTGCCATTTGCCCTGT	1489
QY	1112	GCAGTTTGAACACAGAAATTCGCTGTGA	1139
Db	1490	CCAGCTTCGACAGCAAGCTGGCGTGCA	1517

RESULT 7

US-09-032-365A-61
; Sequence 61, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502th, Michael
; APPLICANT: Nishina, Patsy
; APPLICANT: Naggart, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto

; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,365A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1890 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-032-365A-61

Query Match 5.9%; Score 66.8; DB 3; Length 1890;
Best Local Similarity 57.1%; Pred. No. 2.2e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY	872	TGCTGCTGAGCAACAAGCTCCACGGTGGCATGACGAGTTACGTTCTGGTCTTAAAT	931
Db	1271	TCGAGCTGCAAAACAGACACCTGCTGGAATGATGACACACAGTCTATGACTCACT	1330
QY	932	TCCATGCTCGAGTCACAGTGGCTTCGGTTAAGAACTTCCAGCTTTGGCAGTTAGTGACT	991
Db	1331	TCCATGGCGCGTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGCGCAATGACC	1390
QY	992	GTGAAGCAGGCGACAGATCTGAGAGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA	1051
Db	1391	-----CGGACTACATCGTGATGCGAGTTTGGCGGCTAGCAGAGGATG	1432
QY	1052	TGTTTACCATGGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGCTATCTGCTGA	1111
Db	1433	TGTTTACCATGGATTATCACTCCCGCTGTGTGCACTGCAGGCTTTGCCATTTGCCCTGT	1492
QY	1112	GCAGTTTGAACACAGAAATTCGCTGTGA	1139
Db	1493	CCAGCTTCGACAGCAAGCTGGCGTGCA	1520

RESULT 8

US-09-032-365A-59
; Sequence 59, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: No. 6114502th, Michael
; APPLICANT: Nishina, Patsy
; APPLICANT: Naggart, Juergen
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto

```
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FatSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,365A
; FILING DATE:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1936 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-032-365A-59

Query Match 5.9%; Score 66.8; DB 3; Length 1936;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGTGTCGTGACACAAAGCTCCACGGTGGCATGAGCAGTTACCTTGTGCTTAAATT 931
Db 1316 TCAGCTGCAAAACACACACCTGTCTGGAATGATGACACAGTCTCTATGTACTCACT 1375
QY 932 TCCATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1376 TCCATGGCGCGTCACACAGCCCTCGTGAAGAACTTTCAGATCATCCATGGCAATGACC 1435
QY 992 GTGAAGCAGGCGACATCTGAGAGGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1436 -----CGGACTACATCGTGATGCGAGTTTGGCCGGTAGCAGAGGATG 1477
QY 1052 TGTTTACCATTGATATGATATCCGATTTCTCGGTTTCAGCGTTTGCTATCTGCTGA 1111
Db 1478 TGTTTACCATTGATATCAACTACCCGCTGTGTGCACTGCGAGGCTTTGCCAATGGCCCTGT 1537
QY 1112 GCAGTTTGAAGAACAGAAATTCGCTGTGA 1139
Db 1538 CCAGTTCGACAGCAAGCTGGCGTGCGA 1565

RESULT 9
US-08-631-200-7
; Sequence 7, Application US/08631200
; Patent No. 5646040
; GENERAL INFORMATION:
; APPLICANT: Kleyn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.

; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/631,200
; FILING DATE: 12-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: DNA (genomic)
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 153..1670
; US-08-631-200-7

Query Match 5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGTGTCGTGACACAAAGCTCCACGGTGGCATGAGCAGTTACCTTGTGCTTAAATT 931
Db 1420 TCAGCTGCAAAACACACACCTGTCTGGAATGATGACACAGTCTCTATGTACTCACT 1479
QY 932 TCCATGTCGAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1480 TCCATGGCGCGTCACACAGCCCTCGTGAAGAACTTTCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAAGCAGGCGACATCTGAGAGGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTGATGCGAGTTTGGCCGGTAGCAGAGGATG 1581
QY 1052 TGTTTACCATTGATATGATATCCGATTTCTCGGTTTCAGCGTTTGCTATCTGCTGA 1111
Db 1582 TGTTTACCATTGATATCAACTACCCGCTGTGTGCACTGCGAGGCTTTGCCAATGGCCCTGT 1641
QY 1112 GCAGTTTGAAGAACAGAAATTCGCTGTGA 1139
Db 1642 CCAGTTCGACAGCAAGCTGGCGTGCGA 1669

RESULT 10
US-08-829-553-7
; Sequence 7, Application US/08829553
; Patent No. 5817762
; GENERAL INFORMATION:
; APPLICANT: Kleyn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/829,553
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/631,200
; FILING DATE: 12-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 153..1670
; US-08-829-553-7

Query Match 5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGCTGGCATGAGCAGTTACGTTGCTGGTGTAAATT 931
Db 1420 TCGAGCTGCACAAACAGACACCTGTCTGAATGATGACACACAGTCTATGACTCAACT 1479

QY 932 TCATGGTTCGAGTCACAGTGGCTTCGGTTAAGAACTTTTCAGCTTTGGCAGTTAGTGACT 991
Db 1480 TCCATGGGCGCTGCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539

QY 992 GTGAAGCAGGGCAGACATCTGAGAGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTATGATGAGTTGGCCGGTAGCAGAGATG 1581

QY 1052 TGTTTACCATGGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGCTATCTGCTGA 1111
Db 1582 TGTTCACCATGGATTACAACTACCCGCTGTGTGCACTGCGAGGCTTTGCCATTGCCCTGT 1641

QY 1112 GCAGTTTGAACACAGAATTGCTGTGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

RESULT 11
US-08-922-267A-7
; Sequence 7, Application US/08922267A
; Patent No. 5861239
; GENERAL INFORMATION:
; APPLICANT: Klevn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/829,553
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/631,200
; FILING DATE: 12-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 153..1670
; US-08-922-267A-7

Query Match 5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGCTGGCATGAGCAGTTACGTTGCTGGTGTAAATT 931
Db 1420 TCGAGCTGCACAAACAGACACCTGTCTGAATGATGACACACAGTCTATGACTCAACT 1479

QY 932 TCATGGTTCGAGTCACAGTGGCTTCGGTTAAGAACTTTTCAGCTTTGGCAGTTAGTGACT 991
Db 1480 TCCATGGGCGCTGCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539

QY 992 GTGAAGCAGGGCAGACATCTGAGAGATCATCTCCAGTTTGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTATGATGAGTTGGCCGGTAGCAGAGATG 1581

QY 1052 TGTTTACCATGGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGCTATCTGCTGA 1111
Db 1582 TGTTCACCATGGATTACAACTACCCGCTGTGTGCACTGCGAGGCTTTGCCATTGCCCTGT 1641

QY 1112 GCAGTTTGAACACAGAATTGCTGTGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

RESULT 12
US-08-936-707A-7
; Sequence 7, Application US/08936707A
; Patent No. 5871931
; GENERAL INFORMATION:
; APPLICANT: Klevn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,707A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 153..1670
; US-08-936-707A-7

Query Match          5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGGTTCAGTTCGTTGGTGTAAAT 931
Db 1420 TCGAGCTGCMAAACAGACACCTGCTGGAATGATGACACACAGTCTATGTACTCAACT 1479

QY 932 TCCATGGTGCAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1480 TCCATGGGCGCTGCACAGAGGCTCCGTTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539

QY 992 GTGAAGCAGGCGACATCTGAGAGGATCATATCTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTGATGAGTTTGGCCGGGTAGCAGAGGATG 1581

QY 1052 TGTTTACCATGGATTATGGATATCCGATTTTCGCGTTTCAAGCGTTTGTCTATCTGCCTGA 1111
Db 1582 TGTTCACCATGGATTACAACTACCGCTGTGTGCACTGCGAGGCTTTGCCATTTGCCCTGT 1641

QY 1112 GCAGTTTTCGAACCAAGAAATTCGCTGGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

RESULT 13
US-08-936-706A-7
; Sequence 7, Application US/08936706A
; Patent No. 5876919
; GENERAL INFORMATION:
; APPLICANT: Kley, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
```

```
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,706A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-099
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 153..1670
; US-08-936-706A-7

Query Match          5.9%; Score 66.8; DB 2; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;

QY 872 TGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCAGGTTCAGTTCGTTGGTGTAAAT 931
Db 1420 TCGAGCTGCMAAACAGACACCTGCTGGAATGATGACACACAGTCTATGTACTCAACT 1479

QY 932 TCCATGGTGCAGTCACAGTGGCTTCGGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
Db 1480 TCCATGGGCGCTGCACAGAGGCTCCGTTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539

QY 992 GTGAAGCAGGCGACATCTGAGAGGATCATATCTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
Db 1540 -----CGGACTACATCGTGATGAGTTTGGCCGGGTAGCAGAGGATG 1581

QY 1052 TGTTTACCATGGATTATGGATATCCGATTTTCGCGTTTCAAGCGTTTGTCTATCTGCCTGA 1111
Db 1582 TGTTCACCATGGATTACAACTACCGCTGTGTGCACTGCGAGGCTTTGCCATTTGCCCTGT 1641

QY 1112 GCAGTTTTCGAACCAAGAAATTCGCTGGA 1139
Db 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

RESULT 14
US-09-248-203-7
; Sequence 7, Application US/09248203
; Patent No. 6043346
; GENERAL INFORMATION:
; APPLICANT: Kley, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
```

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/248,203
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/936,707
FILING DATE: 24-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2040 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 153..1670
US-09-248-203-7

Query Match 5.9%; Score 66.8; DB 3; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;
QY 872 TGGTCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTCAGTTCGCTGGTCTTAAATT 931
DB 1420 TCAGCTGCAAAACAAGACACCTGTCTGGAATGATGACACACAGTCCTATGTACTCACT 1479
QY 932 TCCATGGTCGAGTCACAGTGGCTTCGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
DB 1480 TCCATGGCGCGTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAAGCAGGGCAGACATCTGAGAGGATCATATCTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
DB 1540 -----CGGACTACATCGTGATGAGTTTGGCCGGGTAGCAGAGATG 1581
QY 1052 TGTTCACCATGGATTATGAGATATCCGATTTTCGCGTTTCAAGGTTTGTCTATCTGCCTGA 1111
DB 1582 TGTTCACCATGGATTATCAACTACCGCTGTGTGCACTGCAGGCTTTGCCATTGCCCTGT 1641
QY 1112 GCAGTTTGAACACCAAGTTGCCCTGGA 1139
DB 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

RESULT 15
US-09-406-071-7
Sequence 7, Application US/09406071
Patent No. 6207386
GENERAL INFORMATION:
APPLICANT: Klevn, Patrick W.
APPLICANT: Moore, Karen J.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF INVENTIONS: 60
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/406,071
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/936,707
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2040 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 153..1670
US-09-406-071-7

Query Match 5.9%; Score 66.8; DB 3; Length 2040;
Best Local Similarity 57.1%; Pred. No. 2.3e-11;
Matches 153; Conservative 0; Mismatches 97; Indels 18; Gaps 1;
QY 872 TGGTCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTCAGTTCGCTGGTCTTAAATT 931
DB 1420 TCAGCTGCAAAACAAGACACCTGTCTGGAATGATGACACACAGTCCTATGTACTCACT 1479
QY 932 TCCATGGTCGAGTCACAGTGGCTTCGTTAAGAACTTTCAGCTTGTGGCAGTTAGTGACT 991
DB 1480 TCCATGGCGCGTCACACAGGCTCCGTGAAGAACTTCCAGATCATCCATGGCAATGACC 1539
QY 992 GTGAAGCAGGGCAGACATCTGAGAGGATCATATCTCCAGTTTGGGAAAGTTGGGAAGGACA 1051
DB 1540 -----CGGACTACATCGTGATGAGTTTGGCCGGGTAGCAGAGATG 1581
QY 1052 TGTTCACCATGGATTATGAGATATCCGATTTTCGCGTTTCAAGGTTTGTCTATCTGCCTGA 1111
DB 1582 TGTTCACCATGGATTATCAACTACCGCTGTGTGCACTGCAGGCTTTGCCATTGCCCTGT 1641
QY 1112 GCAGTTTGAACACCAAGTTGCCCTGGA 1139
DB 1642 CCAGCTTCGACAGCAAGCTGGCGTGCGA 1669

Search completed: December 26, 2005, 06:53:57
Job time : 269 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 03:41:26 ; Search time 654 Seconds
(without alignments)
904.600 Million cell updates/sec

Title: US-10-763-042-20

Perfect score: 1140

Sequence: 1 atgagctccgaagtact.....aaaccagaattgctgtgaa 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4168288 seqs, 259477437 residues

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:
- 2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:
- 3: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:
- 4: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:
- 5: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
- 6: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:
- 7: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:
- 8: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq2:
- 9: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq3:
- 10: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	53	4.6	924	6	US-10-750-185-31303
C 2	36	3.2	1293	6	US-10-525-710-7
C 3	35.6	3.1	100387	7	US-11-121-086-94
C 4	34.8	3.1	1161	6	US-10-667-295-179
C 5	33.8	3.0	1389	6	US-10-750-185-64773
C 6	33.8	3.0	110711	6	US-10-995-561-13264
C 7	33.6	2.9	315	6	US-10-802-796-127
C 8	33	2.9	2848	6	US-10-750-185-56373
C 9	32.6	2.9	867	6	US-10-750-185-26853
C 10	32	2.8	2195	6	US-10-750-185-41445
C 11	31.8	2.8	1172	6	US-10-750-185-55303
C 12	31.8	2.8	1904	6	US-10-750-185-57408
C 13	31.4	2.8	600	6	US-10-750-185-2925
C 14	31.4	2.8	783	6	US-10-750-185-57709
C 15	31.2	2.7	1587	6	US-10-750-185-25701
C 16	31.2	2.7	4608	6	US-10-821-234-137
C 17	31.2	2.7	151169	7	US-11-121-086-38
C 18	30.6	2.7	909	6	US-10-750-185-41557
C 19	30.6	2.7	1499	6	US-10-750-185-30575
C 20	30.6	2.7	3234	6	US-10-750-185-32891
C 21	30.6	2.7	195998	6	US-10-995-561-13489
C 22	30.6	2.7	212716	7	US-11-121-086-95
C 23	30.4	2.7	1015	6	US-10-750-185-39157

C	24	30.4	2.7	1536	6	US-10-750-185-50389	Sequence 50389, A
C	25	30.4	2.7	207600	7	US-11-112-908-31	Sequence 21, Appl
C	26	30	2.6	1849	6	US-10-750-185-28537	Sequence 28537, A
C	27	30	2.6	2594	6	US-10-750-185-63480	Sequence 63480, A
C	28	29.8	2.6	1207	6	US-10-750-185-40341	Sequence 40341, A
C	29	29.8	2.6	2967	6	US-10-750-185-33220	Sequence 33220, A
C	30	29.8	2.6	156544	7	US-11-121-086-81	Sequence 81, Appl
C	31	29.8	2.6	244196	6	US-10-995-561-13327	Sequence 13327, A
C	32	29.8	2.6	285300	6	US-10-857-780-6	Sequence 6, Appl
C	33	29.6	2.6	1482	7	US-11-137-465-13	Sequence 13, Appl
C	34	29.6	2.6	1620	6	US-10-750-185-53467	Sequence 53467, A
C	35	29.6	2.6	4104	6	US-10-131-826A-449	Sequence 449, App
C	36	29.4	2.6	2689	6	US-10-750-185-53445	Sequence 53445, A
C	37	29.4	2.6	157230	7	US-11-112-908-64	Sequence 64, Appl
C	38	29.4	2.6	170508	7	US-11-112-908-62	Sequence 62, Appl
C	39	29.4	2.6	173115	7	US-11-112-908-65	Sequence 65, Appl
C	40	29.4	2.6	182190	7	US-11-121-086-102	Sequence 102, App
C	41	29.4	2.6	185393	7	US-11-121-086-101	Sequence 101, App
C	42	29.4	2.6	186442	7	US-11-121-086-104	Sequence 104, App
C	43	29.2	2.6	378	6	US-10-793-626-1135	Sequence 1135, Ap
C	44	29.2	2.6	378	6	US-10-793-626-2657	Sequence 2657, Ap
C	45	29.2	2.6	931	6	US-10-750-185-30913	Sequence 30913, A

ALIGNMENTS

RESULT 1

US-10-750-185-31303/c
; Sequence 31303, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31303
; LENGTH: 924
; TYPE: DNA
; ORGANISM: Bovine 19866881111971
US-10-750-185-31303

Query Match 4.6%; Score 53; DB 6; Length 924;
Best Local Similarity 62.4%; Mismatches 50; Indels 0; Gaps 0;
Matches 83; Conservative 0

QY	1007	CATCTGAGAGGATCATCTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGGATT	1066
DB	785	CAGCGGACTACTCTGCTCGATTTGGTCGTGTAGCCCGACACACGTTCCACATGACT	726
QY	1067	ATGGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGCTGACGAGTTTGAACCA	1126
DB	725	TCGCGTTCCACTTTGCCCGCTCCAGCCTTTGCCATCTGCTTGTCCAGTTTCGATGGGA	666
QY	1127	GAATGCTCTGCA	1139
DB	665	AGCTGCGATGCA	653

RESULT 2

US-10-525-710-7
; Sequence 7, Application US/10525710

Qy	971	AGCTTTGGCAGTTAGTGACTGTGAAGCAGGGGAGACATCTTGAGAGGATCATACTCCAGT	1033
Db	82657	GTCTTCTCTGGTGTGTTTGTGACTACTGCGTCCCAAAATACTCTCTCAAGATACATTTCTTAAT	82595
Qy	1031	TTGGGAAAGTTGGGAAGGACATGTTTACCATGATTATGGATATCCGATTTCTTCGCGTTTC	1090
Db	82597	TGCTTAATTTGACCAATTAATCTCTGTTTAAATCTTGTCTATATATGTTTCTCTCTTAT	82533
Qy	1091	AAGCGTTTGGCTATC 1104	
Db	82537	TAGCTTTCCAGATC 82524	
RESULT 4			
US-10-667-295-179			
; Sequence 179, Application US/10667295			
; Publication No. US20050257293A1			
; GENERAL INFORMATION:			
; APPLICANT: Mascia, Peter			
; TITLE OF INVENTION: BIOLOGICAL CONTAINMENT SYSTEM			
; FILE REFERENCE: 11696-047001			
; CURRENT APPLICATION NUMBER: US/10/667,295			
; CURRENT FILING DATE: 2003-09-17			
; PRIOR APPLICATION NUMBER: US 60/411,823			
; PRIOR FILING DATE: 2002-09-17			
; NUMBER OF SEQ ID NOS: 263			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 179			
; LENGTH: 1161			
; TYPE: DNA			
; ORGANISM: Triticum aestivum			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1)-(1161)			
; OTHER INFORMATION: Ceres Seq. ID no. 12598265			
US-10-667-295-179			
Query Match 3.1%; Score 34.8; DB 6; Length 1161;			
Best Local Similarity 53.7%; Pred. No. 0.24;			
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;			
Qy	968	TTGAGCTTTGGCGAGTTAGTGACTGTGAAGCGGCGAGCATCTGAGAGATCATACTCC	1027
Db	926	TTGAGCGTGTTCACCAAGCTAACCGTTAAGATGGATGTGATCTGAGGGTCTCATGCTTT	985
Qy	1028	AGTTTGGGAAGTTGGGAAGGACATGTTTACCATGGATTATGGATATCCGATTTCTGCGT	1087
Db	986	TGGCTGGAGATTTTGGACTGAATGTGTTTACAGATATCTAGAACTGTTTAGCTATGC	1045
Qy	1088	TTCAAGCGTTTGCT 1101	
Db	1046	TTTAGTTTGTGAT 1059	
RESULT 5			
US-10-750-185-64773/c			
; Sequence 64773, Application US/10750185			
; Publication No. US20050260603A1			
; GENERAL INFORMATION:			
; APPLICANT: MMI GENOMICS, INC.			
; APPLICANT: DENISE, Sue K.			
; APPLICANT: KERR, Richard			
; APPLICANT: ROSENFELD, David			
; APPLICANT: HOLM, Tom			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10/750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; NUMBER OF SEQ ID NOS: 64922			

```
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 64773
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Bovine 19866880552958
US-10-750-185-64773

Query Match          3.0%; Score 33.8; DB 6; Length 1389;
Best Local Similarity 48.7%; Pred. No. 0.6;
Matches 92; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 201 CGTTTCCTAGTCGAGGATTCACCAAGGAGATTGTAGCTGTTCCTCGAATTCCTC 260
DB 1046 CTTTGAAGAGCTGATGCTAGTATAGAGAAATGTATCTGATAGATAATGCTCAA 987

QY 261 TAAATGACTTTCCCTATCTCCCTCAAGCAGTCTGTCGAAGAGATTCCTAGTCCAATG 320
DB 986 GAAGATTACTGACTATTGCTCTCTCGCAGCAAAATATCAAGAAATTTTGGTGACATGG 927

QY 321 CTTTATAAAGCTAATCGAATACTCAATCGTATCATCTCTATCTCGGATTAACTACCTC 380
DB 926 TTTAAGGAATAATCTTTAAATACTACCAGCATGCACACATCTTAGACATATGTTCCAG 867

QY 381 TTTGACGGA 389
DB 866 GTCCATGGA 858

RESULT 6
US-10-995-561-13264/c
; Sequence 13264, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13264
; LENGTH: 110711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(110711)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13264

Query Match          3.0%; Score 33.8; DB 6; Length 110711;
Best Local Similarity 54.4%; Pred. No. 10;
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 259 TCTAATTCGACTTTCCTATCTCCCTCAAGCAGTCTGTCGAAGAGATTCCTAGTTCAA 318
DB 66003 TTTAAATTTGTTTCAGAGATTCCTCCCTACCTTTAGTCATAAAGATACTCTCCTTTAT 65944

QY 319 TCGTTTATAAAGCTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 378
DB 65943 TGCCTTCTATTAACTAAGAGTTATCTCTCACCTTAGATCTTTAAACCCATCCAGAGTCCA 65884

QY 379 TCTTT 383
DB 65883 CCTTT 65879

RESULT 7
US-10-802-796-127
; Sequence 127, Application US/10802796
; Publication No. US20050250104A1
```

```
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; APPLICANT: BUCHRIESER-BROSCH, ROLAND
; APPLICANT: GORDON, STEPHEN
; APPLICANT: BILLAULT, ALAIN
; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
; TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: MYCOBACTERIA.
; FILE REFERENCE: 05394.0011-00000
; CURRENT APPLICATION NUMBER: US/10/802,796
; CURRENT FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: US/09/673,476
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/IB99/00740
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 09/060,756
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIN Ver. 2.2
; SEQ ID NO 127
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (129)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (153)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (231)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (258)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (298)
; OTHER INFORMATION: a, t, c or g
US-10-802-796-127

Query Match          2.9%; Score 33.6; DB 6; Length 315;
Best Local Similarity 62.2%; Pred. No. 0.27;
Matches 51; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 105 GTCCCGGAGGAGCTCTTAGAGAAATCTGATAGGTTGAGACTGTTCAGCGGCGCGA 164
DB 214 GCGCGGAGGAGTCTGTCNTTGGATCCCGGCTGGCATTCGCGTGTGGCGGCGCGCG 273

QY 165 TTGCGCGCTCGCGCGCAACGCTG 186
DB 274 GTGGTGGGGGGGCAACGCTG 295

RESULT 8
US-10-750-185-56373/c
; Sequence 56373, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
```


Db 850 TTTTCTTTTAACTTGGAGGAGCTATATAATATATGATCATATACCATTTCTTTCATAT 791
QY 369 ATTAACACTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTTCTTAAGCTGAAGCG 428
Db 790 TTTTTCACAAAACCTTTGGTAGCATATAATAAATTAATCTGACAGCTTCAATGCCAAAATG 731
QY 429 CGCAACTTGCACCTGATTAC 447
Db 730 TGAATTTTCTACTTAAAC 712

RESULT 12

US-10-750-185-57408/c
; Sequence 57408, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57408
; LENGTH: 1904
; TYPE: DNA
; ORGANISM: Bovine 19866881175911
US-10-750-185-57408

Query Match 2.8%; Score 31.8; DB 6; Length 1904;
Best Local Similarity 48.6%; Pred. No. 3.6;
Matches 87; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 205 TGTCTAGCTGGAGGATTCACCAAGGAGATTGTAGCTGTTCTCTGAAATCTCTCTTAA 264
Db 656 TGGACTAGCTGGTTCAATTCCTCTATCCGTTGTTATTGTAGCTTAGGTCCTAGGTTCA 597
QY 265 TTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCAGAGATTCTCTAGTTCAATGCTTT 324
Db 596 AAACCATGTCAATCACACATTTGGAGCAGAGAGAGGATACCCCATCACAAAGTACA 537
QY 325 ATAAACGTAATCGAAATATCTCAATCGTATCATCTCTATCTCGGATTAACCTCTTT 383
Db 536 TTCAGCTGAGCAGCTTTTCAATACTTTTATATTTTCTTGAATTAATTAGAGCCTT 478

RESULT 13

US-10-750-185-2925/c
; Sequence 2925, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2925
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT17358
US-10-750-185-2925

Query Match 2.8%; Score 31.4; DB 6; Length 600;
Best Local Similarity 48.0%; Pred. No. 2.3;
Matches 86; Conservative 1; Mismatches 92; Indels 0; Gaps 0;

QY 205 TGTCTAGCTGGAGGATTCACCAAGGAGATTGTAGCTGTTCTCTGAAATCTCTCTTAA 264
Db 393 TGGACTAGCTGGTTCAATTCCTCTATCCGTTGTTATTGTAGCTTAGGTCCTAGGTTCA 334
QY 265 TTGACTTTCCCTATCTCCCTCAAGCAGTCTGGTCCAGAGATTCTCTAGTTCAATGCTTT 324
Db 333 AAACCATGTCAATCACACATTTGGAGCAGAGAGGATACCCCATCACAAAGTACA 274
QY 325 ATAAACGTAATCGAAATATCTCAATCGTATCATCTCTATCTCGGATTAACCTCTTT 383
Db 273 TTCAGCTGAGCAGCTTTTCAATACTTTTATATTTTCTTGAATTAATTAGAGCCTT 215

RESULT 14

US-10-750-185-57709/c
; Sequence 57709, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57709
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Bovine 19866880388936
US-10-750-185-57709

Query Match 2.8%; Score 31.4; DB 6; Length 783;
Best Local Similarity 59.6%; Pred. No. 2.7;
Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 363 TCTCGATTAACTACTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTTCTAAGCT 422
Db 719 TTTCCAGTTTCCCTCACCCCTATTACCCATAATGCTGTTCTCTCTGCTTGTCTCTGCT 660
QY 423 GAAGCGCAACTTGCACCTGATTACATCA 451
Db 659 GCAGAGTGGGAGCGGCACTGCTTTCAACA 631

RESULT 15

US-10-750-185-25701/c
; Sequence 25701, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David

```

; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 25701
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-25701

```

	Query Match	2.7%;	Score 31.2;	DB 6;	Length 1587;
	Best Local Similarity	54.3%;	Pred. No. 5.1;		
	Matches	63;	Conservative	0;	Mismatches 53; Indels 0; Gaps 0;
Qy	341	ATACTCAATCGTATCATCTCTATCTCGATTAACTACCTCTTTGACGGATAACGGGAAGT	400		
Db	663	ACACTCCATGTATCAGGATTCAAGGGAACATGACACCTCTGTTCAGAAAAACAAAAAGG	604		
Qy	401	TTCTTTCTTGCTGCTTCTTAAGCTGAAGCGCGCACTTGCATGATTACATCATCTCT	456		
Db	603	ATCTTGTGCAGAGATTTCAATTGAATTTCTCACCTGGCATCTCCACATTAGCACT	548		

Search completed: December 26, 2005, 07:23:33
Job time : 657 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2005, 22:15:46 ; Search time 796 Seconds
(without alignments)
9544.912 Million cell updates/sec

Title: US-10-763-042-20
Perfect score: 1140
Sequence: 1 atgaggtccgaagttact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1140	100.0	1140	14	Adw64874 Arabidops
2	1140	100.0	1143	12	Adn73048 Thale cre
3	572.8	50.2	1419	12	Ado61994 Transcript
4	571.2	50.1	1295	3	Aac37041 Arabidops
5	567.8	49.8	1140	14	Adw64876 Arabidops
6	396	34.7	396	8	Abx62764 Arabidops
7	387.4	34.0	1681	14	Adw16693 Eucalyptu
8	367.8	32.3	1629	13	Adx13792 Plant ful
9	354.8	31.1	2586	14	Adw16694 Eucalyptu
10	342.6	30.1	1218	14	Adw64868 Arabidops
11	342.6	30.1	1221	12	Adn73404 Thale cre
12	336.2	29.5	1948	14	Aeb67645 Rice geno
13	326	28.6	469	6	Ab193451 Arabidops
14	324.6	28.5	1887	14	Adw17065 Pinus rad
15	322	28.2	1923	14	Adw17066 Pinus rad
16	321	28.2	2181	14	Adw17064 Pinus rad
17	282.8	24.8	2444	14	Adw17062 Pinus rad
18	254	22.3	1911	14	Adw18185 Pinus rad
19	254	22.3	1941	14	Adw18186 Pinus rad

20	250.4	22.0	1182	14	Adw64867 Arabidops
21	248.8	21.8	1501	14	Adw17063 Pinus rad
22	245.2	21.5	1991	14	Adw16695 Eucalyptu
23	229.2	20.1	1002	13	Adx36491 Plant ful
24	210.2	18.4	1164	14	Adw64871 Arabidops
25	195.4	17.1	1335	14	Adw64875 Arabidops
26	195.4	17.1	1338	10	Adx37190 Plant yie
27	195.4	17.1	1338	12	Adi44432 Plant tra
28	193.6	17.0	1792	14	Adw16699 Eucalyptu
29	193.6	17.0	1852	14	Adw16697 Eucalyptu
30	193	16.9	1137	14	Adw64872 Arabidops
31	192	16.8	1980	14	Adw16698 Eucalyptu
32	188.2	16.5	1380	3	Aac33976 Arabidops
33	184	16.1	1287	14	Adw64870 Arabidops
34	184	16.1	1567	12	Ado61990 Transcript
35	169.8	14.9	1910	13	Adx47926 Plant ful
36	169.8	14.9	1930	13	Adx48044 Plant ful
37	169.6	14.5	1822	14	Adw16696 Eucalyptu
38	164.8	14.5	1414	13	Adx54648 Plant ful
39	162	14.2	1926	13	Adx51972 Plant ful
40	160.8	14.1	1750	13	Adx59298 Plant ful
41	160.4	14.1	571	13	Acn49879 Cotton pr
42	160.4	14.1	1030	13	Adr60407 Cotton cd
43	159.2	14.0	1800	14	Aeb67826 Rice geno
44	148.4	13.0	1365	14	Adw64866 Arabidops
45	146.4	12.8	1599	13	Adx47638 Plant ful

ALIGNMENTS

RESULT 1
ADW64874
ID ADW64874 standard; cDNA; 1140 BP.
XX
AC ADW64874;
XX
DT 07-APR-2005 (first entry)

DE Arabidopsis thaliana TUBBY-like protein (TLP) 9 encoding cDNA.

KW Transgenic plant; salt tolerance; crop improvement; cold tolerance;
KW oxidative stress; drought resistance; TUBBY-like protein; plant;
KW chromosome 2; gene; ss.

OS Arabidopsis thaliana.

Key Location/Qualifiers

CDS 1..1140

FT /*tag= a

FT /product= "Arabidopsis thaliana TLP9 protein"

FT /partial

FT /note= "No stop codon"

XX US2005014266-A1.

XX 20-JAN-2005.

XX 21-JAN-2004; 2004US-00763042.

XX 21-JAN-2003; 2003US-0441380P.

XX (SINT-) ACAD SINICA.

XX Shaw J, Lai C;

XX WPI; 2005-080955/09.

XX P-PSDB; ADW64863.

XX GENBANK; AF487270.

PT New isolated Arabidopsis TUBBY-like proteins, useful for producing transgenic or transformed cells or plants having higher tolerance to salt, chilling, pathogens, oxidative stress, or water-deficit.

CC sequences can also be useful as positive or negative selectable markers
 CC during transformation of cells or tissues. The identified genes play a
 CC role in a variety of biological processes such as DNA replication, cell
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
 CC transcription factors. This polynucleotide sequence is the cress cDNA
 CC upregulated 1.3 fold or more in plants overexpressing the E2Fa/Dpa
 CC transcription factor, given in an exemplification of the invention.
 XX
 SQ Sequence 1143 BP; 287 A; 265 C; 279 G; 312 T; 0 U; 0 Other;

Query Match 100.0%; Score 1140; DB 12; Length 1143;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGTTCCGAGATTACTCCAGGAAATGCGGTCTAGGCCAACCGTGTAGTTTCAGGC 60
 DB 1 ATGACGTTCCGAGATTACTCCAGGAAATGCGGTCTAGGCCAACCGTGTAGTTTCAGGC 60

QY 61 GCGGCTCAACCGCTAATAGTTTCCAGACCTTTCAGCTGTCGAGCTCCGCGAGGAGCTG 120
 DB 61 GCGGCTCAACCGCTAATAGTTTCCAGACCTTTCAGCTGTCGAGCTCCGCGAGGAGCTG 120

QY 121 CTTAGAGAAATCTGATTAGGTTGAGACTGTTGACGGCGGATTTGGCGTCCGCGCGCA 180
 DB 121 CTTAGAGAAATCTGATTAGGTTGAGACTGTTGACGGCGGATTTGGCGTCCGCGCGCA 180

QY 181 AACGTGGTGGTTCGCGCGCTTTGCTGAGTGTGAGGATTTCTCAAGGAGATTGTA 240
 DB 181 AACGTGGTGGTTCGCGCGCTTTGCTGAGTGTGAGGATTTCTCAAGGAGATTGTA 240

QY 241 GCTGTTCTCGAATTCCTCTAAATGACTTTCCTATCCCTCAAGCAGTCTGTTCCA 300
 DB 241 GCTGTTCTCGAATTCCTCTAAATGACTTTCCTATCCCTCAAGCAGTCTGTTCCA 300

QY 301 AGAGATTCTCTAGTTCAATGCTTTATAAAGCTTAATCGAAATCTCAATCGTATCATCTC 360
 DB 301 AGAGATTCTCTAGTTCAATGCTTTATAAAGCTTAATCGAAATCTCAATCGTATCATCTC 360

QY 361 TATCTCGAATTAATACCTCTTTTGACGATTAACGGGAAGTTTCTTCTGCTGTTCTAAG 420
 DB 361 TATCTCGAATTAATACCTCTTTTGACGATTAACGGGAAGTTTCTTCTGCTGTTCTAAG 420

QY 421 CTGAGCGCGCACTTGCATGATTAATCATCTCTTTTGGTTTCAGACGATATCTCAAG 480
 DB 421 CTGAGCGCGCACTTGCATGATTAATCATCTCTTTTGGTTTCAGACGATATCTCAAG 480

QY 481 AGAAGCAACCGGTATCTGCGGAATGAGATCGAACTTCTTGGAAACAAATTCAGGTC 540
 DB 481 AGAAGCAACCGGTATCTGCGGAATGAGATCGAACTTCTTGGAAACAAATTCAGGTC 540

QY 541 TTTGATGTTAGTCAGACCGGAGCAGGAAAGATGACAGAGCGGCTCTTCTAATTTTCATC 600
 DB 541 TTTGATGTTAGTCAGACCGGAGCAGGAAAGATGACAGAGCGGCTCTTCTAATTTTCATC 600

QY 601 AAAGTTTCACTAGATTCTTCHAGGAAGTTTACCCATGCTCACAATTTCAACGAGTTA 660
 DB 601 AAAGTTTCACTAGATTCTTCHAGGAAGTTTACCCATGCTCACAATTTCAACGAGTTA 660

QY 661 AACGCTTAGGCTCTCGGGACGAGAAAGTGTGATCATGATGATCAATACCTATG 720
 DB 661 AACGCTTAGGCTCTCGGGACGAGAAAGTGTGATCATGATGATCAATACCTATG 720

QY 721 AGCATCGTGGAGTTCGCGAGGATGATGATTTCAACATCCATAGCTCTTTTCCAGTCGG 780
 DB 721 AGCATCGTGGAGTTCGCGAGGATGATGATTTCAACATCCATAGCTCTTTTCCAGTCGG 780

QY 781 TCATCACCAAGTCTTTAGGTTCTCACTCAAAACCAATTCGCGAGTAATAGTGCATATGAC 840
 DB 781 TCATCACCAAGTCTTTAGGTTCTCACTCAAAACCAATTCGCGAGTAATAGTGCATATGAC 840

QY 841 GACTCAGGCAACACTGCGAGATCCACATTTGTTGCTGAGCAACAAAGTCCAGTCGG 900
 DB 841 GACTCAGGCAACACTGCGAGATCCACATTTGTTGCTGAGCAACAAAGTCCAGTCGG 900

QY 901 CATGACAGTTAGCTTGTCTGGTCTTAAATTTCCATGTCGAGTCACAGTGGCTTCGGTT 960
 DB 901 CATGACAGTTAGCTTGTCTGGTCTTAAATTTCCATGTCGAGTCACAGTGGCTTCGGTT 960

QY 961 AAGAACTTTTCAGCTTGTGCGAGTTAGTGAAGCAGGGGACACATCTGAGAGGATC 1020
 DB 961 AAGAACTTTTCAGCTTGTGCGAGTTAGTGAAGCAGGGGACACATCTGAGAGGATC 1020

QY 1021 ATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATATGGAATCCGATT 1080
 DB 1021 ATACTCCAGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATATGGAATCCGATT 1080

QY 1081 TCTGCGTTTCAAGCGTTTCTCTATCTGCTGAGCAGTTTGTGAACACGAAATTCCTGTGA 1140
 DB 1081 TCTGCGTTTCAAGCGTTTCTCTATCTGCTGAGCAGTTTGTGAACACGAAATTCCTGTGA 1140

RESULT 3
 ADO61994
 ID ADO61994 standard; DNA; 1419 BP.
 XX
 AC ADO61994;
 DT 15-JUL-2004 (first entry)
 XX
 DE Transcription factor G2603 coding sequence, SEQ ID 461.
 XX
 KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
 KW osmotic stress tolerance; cold tolerance; heat tolerance;
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyphosate resistance; flowering; fertility; seed development; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2004031349-A2.
 XX
 PD 15-APR-2004.
 XX
 PF 18-SEP-2003; 2003WO-US030292.
 XX
 PR 18-SEP-2002; 2002US-0411837P.
 PR 17-DEC-2002; 2002US-0434166P.
 PR 24-APR-2003; 2003US-0465809P.
 XX
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX
 XX Jiang C. Heard JE. Ratcliffe O. Creelman RA, Adam LJ, Reuber TL;
 PI Riechmann JL, Haake V, Dubell AN, Keddie JS, Sherman BK;
 PI
 XX WPI; 2004-330163/30.
 DR P-PSDB; ADO61995.
 DR
 XX New recombinant polynucleotide encoding transcription factor
 PT polypeptides, useful for producing transgenic plants with advantageous
 PT properties compared to a reference plant.
 XX
 PS Claim 1; SEQ ID NO 461; 510pp; English.
 PS
 XX The present invention relates to novel plant transcription factor
 CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
 CC sequences can be used to produce transgenic plants, which overexpress
 CC (II), where the transgenic plant has an altered trait as compared to a
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises
 CC an altered trait selected from increased tolerance to abiotic stress,
 CC increased tolerance to osmotic stress, increased tolerance to cold,
 CC increased germination in cold, increased tolerance to heat, increased
 CC germination in heat, increased tolerance to freezing conditions,
 CC increased tolerance to low nitrogen conditions, increased tolerance to
 CC low phosphate conditions, increased tolerance to disease, including
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
 CC increased tolerance to multiple fungal pathogens, increased resistance to
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,

CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1419 BP; 361 A; 317 C; 331 G; 410 T; 0 U; 0 Other;

Query Match 50.2%; Score 572.8; DB 12; Length 1419;
Best Local Similarity 71.7%; Pred. No. 1.6e-177;
Matches 852; Conservative 0; Mismatches 267; Indels 69; Gaps 5;

QY 1 ATGACGTCGGAAGTTACTCCAGGAATCGCGTCTAGGCGCACCGTGTAGTTCACGCG 60
DB 66 ATGACCTTACGTAGCTTAATCTTGTAGATGCGTTCGAGACCGCATCGTGTGTCACGAC 125
QY 61 -----GCCGCGCTCAACCGCTAATAGTTTCAGACCGCTTTCAGCTGGTGC 102
DB 126 CTTGCGCGCGCGCAGCTCGCATTCCTTCTGTGTCATCGCAAGATTATCGTGTCA 185
QY 103 GAGCTCCCGGAGAGCTGCTTGAAGAAATCCTGTATTAGGGTTGAGACTGTTGACCGCGGC 162
DB 186 GAGATTCTCGAAGAGCTTCTTAGGAGATTCGTGATTCTGTGTTGAAGCGCGGACGCTGCG 245
QY 163 GATTGCGCTCGCGGAAACGCTGCTGTGCGCGCGGTTTGTCTAGCTGGAGATT 222
DB 246 GGATGCGCGCTCAGCAGCGAGCTGGTGGCTTGTGCGCGGCTTGTCTGGCTGGCGGCTA 305
QY 223 CTCACCAAGGAGATTGTAGCTGTCTGAATTCCTCTAAATTGCACTTTCCTATCTCC 282
DB 306 CTTATGAAAGAACCTGCTGTGCTGAGATCTCTTAAGTTGACTTTCCTCATCTCT 365
QY 283 CTCACGAGCTGGTCCAGAGATTCTCTAGTTCAATGCTTTTATAAAAGTAAATCGAAAT 342
DB 366 CTCACGAGCGCTGCTCCAGGAGTTCACCTGGTTCAATGCTTTTATCAAGAGTAAATCGAAT 425
QY 343 ACTCATCTGATCATCTCTATCTCGGTTAACTACCTCTTTCGCGGATTAACGCGAGTTT 402
DB 426 AGCAATCATATCATCTCTATCTCGGATTAAACCACTCTTTAACGGATGATGGGAATTT 485
QY 403 CTTCTTGTGCTTCTAAGCTGAAGCGCAACTTTCGACTGATTACATCATCTCTTTCGCT 462
DB 486 TTGCTTGTGCTGCTAAGTTGAGACACACACTTGTACGGATTACATTATCTTTACGT 545
QY 463 TCAGACGATATCTCAAGAGAACGCAACGCGTATCTTGGGAGAAATGAGATCGAACTTCCTT 522
DB 546 TCTGATGATATGTCGAGAGAACGCAAGCTTATGTTGGCAAGTGGATCGAACTTCCTA 605
QY 523 GGAACAAATTCAGCTCTTGTAGTGTAGTCAG-----ACCGGACGCGAGAGT 573
DB 606 GGAACGAAATTCATCTGTCTTGTAGTGAATCTGCTGCTTTCACCGGAGCGGAAAGTTG 665
QY 574 CAGAAGCGCGCTCTTCTTAATTTTCATCAAAAGTTTCACCTAGAGTCTCTCAGGGAAGTTAC 633
DB 666 AGAAGAGCGGATCTTATATATCCGCAAGATTTTCAGCAAAAGTTCTCTTGGAGTTAT 725

QY 634 CCCATCGCTCACATTTTCATACGAGTTAAACGCTCTTAGGCTCTCGGGGACCGAGAAAGT 693
DB 726 CTTGCTGCTCATATACATATAGCTGTGATGCTCTTAGGATCCCGGGACCAAGAAAGT 785
QY 694 CGTTGCATCATGATACAAATACCTATATGAGCATCGTGGAGTCGCGAGGAGTAGTAGCTTCA 753
DB 786 CAATGCTTTATGGACACAAATACCTACAGGACCAATGGAGCCTCAAGGAGTAGTAGCTTCA 845
QY 754 ACATCATAGCTCTTTTTCAGTCGCTCATCAACAGTCTTTAGTCTCTCACTCAAAACCA 813
DB 846 CCATCAGAGTTTCCCTTACTCGGTCACTCGGTCAACCTTTATCCAGGTCTCAGTCAAAACCA 905
QY 814 TTGGCGCAGTAAATAGTCATCATGTCAGTCTCAGGCAACCACTCGGAGATCCACCATG 873
DB 906 TTACGAGTA-----GCTCAAGCCACCTGAAGAAACACCATTA 944
QY 874 GTGCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTACGTTGCTGTGCTTAAATTC 933
DB 945 GTGCTGAGCAACAGACACACCGGTGGCAGCAGCAGCTACGCTGCTGTGCTTGAATTC 1004
QY 934 CATGCTCAGTACAGTGGCTTCGTTAAGAACTTTTACGTTTGTGCTGCGAG-----TT 984
DB 1005 CATGGCCGTGTACAGTAGCTGCTGAGGAACCTTTTCAGCTCGTGGCAGGAGCTAGC 1064
QY 985 AGTGACTCTGAAGCAGGCGACACATCTCAGA-----GGATCATACTCCAGTTT 1032
DB 1065 TGTGCGAGTGGCAGCGGAATGTACCGGAGAGGAGCGGAGCGGATTAATGCGATTT 1124
QY 1033 GGGAAAGTTGGGAAGCAGCATGTTTACCATGGATTATGATATCCGATTTCTGCGTTTCAA 1092
DB 1125 GGGAAAGTCGGGAAGATATGTTACGATGATTAATGATACCGATCTCAGCTTTCCAG 1184
QY 1093 GCGTTTGTATCTGCTCAGCAGGTTTGAACCAGAAATTCCTGTGAA 1140
DB 1185 GCTTTTGCCATTTGCTTGAGCAGCTTTGAGACTAGACTGCTTTGTGAA 1232

RESULT 4
AAC37041
ID AAC37041 standard; DNA; 1295 BP.
XX AAC37041;
AC AAC37041;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 15967.
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15967.
XX Hybridization assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-01334256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134321P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136352P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143442P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.

PT New isolated Arabidopsis TUBBY-like proteins, useful for producing
 PT transgenic or transformed cells or plants having higher tolerance to
 PT salt, chilling, pathogens, oxidative stress, or water-deficit.
 XX
 XX
 XX Claim 7; SEQ ID NO 22; 43pp; English.

XX The invention relates to Arabidopsis TUBBY-like proteins (TLP) and their
 CC corresponding nucleic acid sequences. The invention also relates to a
 CC transformed cell or a transgenic plant containing TLP nucleic acid and a
 CC polypeptide and nucleic acids are useful for producing transgenic or
 CC transformed cells or plants having higher tolerance to salt, chilling,
 CC pathogens, oxidative stress or water-deficit. The present sequence is the
 CC Arabidopsis thaliana TLP11 encoding cDNA. The Arabidopsis thaliana TLP11
 CC gene is located at chromosome III.

XX Sequence 1140 BP; 287 A; 268 C; 288 G; 297 T; 0 U; 0 Other;

Query Match 49.8%; Score 567.8; DB 14; Length 1140;
 Best Local Similarity 72.8%; Pred. No. 6.3e-176; Indels 51; Gaps 4;
 Matches 810; Conservative 0; Mismatches 252;
 QY 58 GCGCGCGCTCAACCGCTAATAGTTGACACCCCTTTTTCAGCTGGTTCGGAGCTCCCGGAGGAG 117
 DB 49 GCTGCGGATTCACCTCTCTGTCATCGCAGATTATCGTGGTTCAGAGATTCTCTGAAGAG 108
 QY 118 CTGCTTAGAGAAATCTCGATTAGGTGAGACTGTTGACGGCGCGGCAATTTGGCGCTCGCGG 177
 DB 109 CTTCTTAGGAGATTCTGATTGCTGTTGAAGCGGCGGAGCGGTGGCGGATGGCGCTCAGCA 168
 QY 178 CGAAGAGTGGTGGTGGCGGCTTTCGCTAGCTGGAGGATTCTCACCAAGGAGATT 237
 DB 169 CGAGCGTGGTGGTGGCGGCTTTCGCTAGCTGGAGGATTCTCACCAAGGAGATT 228
 QY 238 GTAGCTGTTCCTGAATTCCTCTTAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGCT 297
 DB 229 GTCGTGTCTCTGAGATCTCTTCTAGTTGACTTTCCCTATCTCTCAAGCAGCTGCT 288
 QY 298 CCAAGAGATTCTCTAGTTCAATGCTTTTATAAAGCTTAATCGAATTAATCAATCGTATCAT 357
 DB 289 CCAAGGATTCACCTGCTTCAATGCTTTTATAAAGCTTAATCGAATTAATCAATCATCAT 348
 QY 358 CTCATCTCTGGATTAACTACTCTTTTGACGGATACGGGAAGTTCTTCTGCTGCTTCT 417
 DB 349 CTCATCTCTGGATTAACTACTCTTTTGACGGATACGGGAAGTTCTTCTGCTGCTTCT 408
 QY 418 AAGCTGAAAGCGGCACTTCTGCTGATTAATCATCTCTTTGCTGCTCAGACGATATCTCA 477
 DB 409 AAGTTGAAGCACAACCTTGTACGGATTACATTAATCTTTAGTTCTGATGATATGTCG 468
 QY 478 AAGAGAGCAACCGGCTATCTTGGGAGAAATGAGATCGAACTTCCTTGGAAACAAATTCACG 537
 DB 469 AGAAGAGCAACCGGCTATCTTGGGAGAAATGAGATCGAACTTCCTTGGAAACAAATTCAC 528
 QY 538 GTCCTTTGATGGTAGTACG-----ACGGAGCAGCGAAGATGACGAAAGAGCGGCTCT 588
 DB 529 GTCCTTTGATGGTAGTACG-----ACGGAGCAGCGAAGATGACGAAAGAGCGGCTCT 588
 QY 589 TCTAATTTTCAAAAGTTTCACTAGATTCTCTCAGGGAAGTTTACCCCATCGCTCAATT 648
 DB 589 TATAATCCCGAAAGTTTTCAGAAAGTTTCTTTGGAAGTTATCTCTGCTGCTATATC 648
 QY 649 TCATACGAGTTAAACGCTCTTAGGCTCTCGGGAGCAGGAAAGATGGTTGCAATCATGGAT 708
 DB 649 ACATATGAGCTGATCTCTTAGGCTCTCGGGAGCAGGAAAGATGCAATGCTCTTATGGAC 708
 QY 709 ACATACCTATGAGCATCTGTGAGTCTCGGAGGAGTAGTAGTTCAATCATCCATAAGCTCT 768
 DB 709 ACAATACCTTACAAGCACAATGGAGCTCAAGGAGTAGTTTCAAGAACCATCAGAGTTTCCC 768
 QY 769 TTTTCAGTGGTTCATCAGCTCTTTAGTCTCTCACTCAAAACCATTTGGCAGTATATCT 828
 DB 769 TTACTCGGTACTCTGGTCAACCTTATCCAGGTCTCTAGTCAAAACCATTTACGCGATA----- 823

QY 829 GCATCATGTAGCGACTCAGGCAACCAACCTGGGAGATCCACCATTGCTGCTGAGCAACAA 888
 DB 824 -----GCTCAAGCCACCTTGAAGAAACCAACCATTAGTGTGCTGAGCAACAG 867
 QY 889 GCTCCACCGTGGCATGAGCAGTTAGCTTGGTGTCTTAAATTTCCATGCTGAGTCACA 948
 DB 868 ACACCAACCGTGGCAGCAGCAGCTACGCTGCTGCTGCTTGAATTTCCATGCTGCTGAC 927
 QY 949 GTGGCTTGGTTAAGAACTTTTTCAGCTTGTGGCAG-----TTAGTGACTGTGAGCA 999
 DB 928 GTAGCGCTCAGTGAAGAACTTTTTCAGCTTGTGGCAGCAGGAGTAGTGTGCTGCTGACG 987
 QY 1000 GGGCAGACATCTGAGA-----GGATCATCTCCAGTTTGGGAAAGTTGGGAAG 1047
 DB 988 GGAATGTCCACCGAGAGGAGCGAGCGGATATATTCAGTTTGGGAAAGTGGGAAA 1047
 QY 1048 GACATGTTTACCATGGAATTCGATATCCGATTTCTGCTGCTTTCAAGCGTTTGTATCTGC 1107
 DB 1048 GATATGTTTACGATGGAATTCGATATCCGATTCAGCTTTCCAGGCTTTTGCCATTGTC 1107
 QY 1108 CTGAGCAGTTTGAACCAACCAATTCCTCTGTGAA 1140
 DB 1108 TTGAGCAGCTTTGAGACTAGAAATCGCTTGTGAA 1140

RESULT 6

ABX62764

ID ABX62764 standard; DNA; 396 BP.

XX AC ABX62764;

XX DT 25-FEB-2003 (first entry)

XX Arabidopsis thaliana expressed sequence related polynucleotide #879.

XX Transgenic plant; plant; genetically modified cell; environmental stress;
 KW ribozyme creation; disease resistance; stress tolerance;
 KW fungicide screening; insecticide screening; gene; ds.

XX Arabidopsis thaliana.

XX US2002040490-A1.

XX PD 04-APR-2002.

XX PF 26-JAN-2001; 2001US-00770423.

XX PR 27-JAN-2000; 2000US-0178512P.

XX (GORL/) GORLACH J.

PA (ANY/) AN Y.

PA (HAMI/) HAMILTON C M.

PA (PRIC/) PRICE J L.

PA (RAIN/) RAINES T M.

PA (YUYI/) YU Y.

PA (RAME/) RAMEAKA J G.

PA (PAGE/) PAGE A.

PA (MATH/) MATHW A V.

PA (LEDF/) LEDFORD B L.

PA (WORS/) WOESSNER J P.

PA (HAAS/) HAAS W D.

PA (GARC/) GARCIA C A.

PA (KRICK/) KRICKER M.

PA (SLAT/) SLATER T.

PA (DAVI/) DAVIS K R.

PA (ALLE/) ALLEN K.

PA (HOFF/) HOFFMAN N.

PA (HURB/) HURBAN P.

XX Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;

PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;

PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;

PI Hurban P;
XX WPI; 2003-110411/10.
XX Novel Arabidopsis thaliana nucleic acid useful for identifying homologous
PT or related genes, and to create genetically modified and transgenic
PT organisms, such as plant cells and plants.
XX
PS Claim 1; SEQ ID NO 879; 43pp; English.
XX
XX The invention describes an Arabidopsis thaliana nucleic acid (I). The
CC polypeptide (II) encoded by (I), transgenic plant (III) or genetically
CC modified cell (IV) are useful for screening a candidate agent for its
CC biological effect, by combining the candidate agent with (II), (III) or
CC (IV), and determining the effect of the candidate agent on (II), (III) or
CC (IV). (I) is useful for identifying homologous or related genes, for
CC producing compositions that modulate the expression or function of its
CC encoded protein, for mapping functional regions of the protein, in
CC diagnosis, for studying associated physiological pathways, for genetic
CC manipulation of cells, preferably plant cells, in screening assays of
CC various plant strains to determine the strains that are capable of
CC withstanding a particular disease or environmental stress, for enhancing
CC or inhibiting production of biosynthetic product in a plant, for
CC producing polypeptides, as probes for the detection of mRNA in biological
CC samples, to generate additional copies of (I), to generate ribozymes or
CC oligonucleotides, as single stranded DNA probes or as triple-strand
CC forming oligonucleotides, and to create genetically modified and
CC transgenic organisms, such as plant cells and plants. (II) or (III) is
CC useful for introducing or improving disease resistance and stress
CC tolerance in plants, screening biological active agents, e.g.,
CC fungicides, insecticides, etc., and for elucidating biochemical pathways.
CC (III) is useful as crops for their enhanced disease resistance, enhanced
CC traits of interest, for screening programs, as crops which exhibit
CC enhanced tolerance to environmental stress, or to produce a factor. This
CC sequence represents a nucleic acid that may correspond to naturally
CC occurring Arabidopsis thaliana expressed sequences. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=999909770423
XX
SQ Sequence 396 BP; 97 A; 91 C; 89 G; 119 T; 0 U; 0 Other;

Query Match 34.7%; Score 396; DB 8; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.7e-119; Indels 0; Gaps 0;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GCGATTGGCGGCGGAAACGTTGGTGGCTTTGTCGGCGGCTTTGTCGTAGCTGGAGG 219
DB 1 GCGGATTGGCGGCTGGCGGAAACGTTGGTGGCTTTGTCGGCGGCTTTGTCGTAGCTGGAGG 60

QY 220 ATTCTCACAAGAGATTGAGCTGTTCTGTAATTCCTCTAAATGACTTTTCCCTATC 279
DB 61 ATTCTCACAAGAGATTGAGCTGTTCTGTAATTCCTCTAAATGACTTTTCCCTATC 120

QY 280 TCCTCAACGAGCTGCTCCAGAGATTCTCTAGTTCAATGCTTTTATAAAGCTATCGA 339
DB 121 TCCTCAACGAGCTGCTCCAGAGATTCTCTAGTTCAATGCTTTTATAAAGCTATCGA 180

QY 340 AATACCTCAATCGTATCATCTCTATCTCGGATTAACCTCTTTGACGGATTAACCGGAG 399
DB 181 AATACCTCAATCGTATCATCTCTATCTCGGATTAACCTCTTTGACGGATTAACCGGAG 240

QY 400 TTTCTTTCTGCTGCTCTTAAGCTGAAGCGCGCAACTTGCACTGATTAATCATCTCTTTTG 459
DB 241 TTTCTTTCTGCTGCTCTTAAGCTGAAGCGCGCAACTTGCACTGATTAATCATCTCTTTTG 300

QY 460 CGTTTCAGACGATATCTCAAGAGAACCAACGCTATCTTTGGGAGATGAGATCGACTTC 519
DB 301 CGTTTCAGACGATATCTCAAGAGAACCAACGCTATCTTTGGGAGATGAGATCGACTTC 360

QY 520 CTTTGGAAACAAAATTCAACGGTCTTTTGATGGTAGTCAG 555
DB 361 CTTTGGAAACAAAATTCAACGGTCTTTTGATGGTAGTCAG 396

RESULT 7

ADW16693
ID ADW16693 standard; cDNA; 1681 BP.
XX
AC ADW16693;
XX
DT 24-MAR-2005 (first entry)
DE
XX Eucalyptus grandis transcription factor cDNA TUBBY family Seq 421.
KW gene; ss; plant; transcription; gene regulation; gene expression;
KW transgenic plant; drought resistance; disease resistance; salt tolerance;
KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
KW flower color.
OS Eucalyptus grandis.
XX
PN WO2005001050-A2.
XX
PD 06-JAN-2005.
XX
PF 07-JUN-2004; 2004WO-US017965.
XX
PR 06-JUN-2003; 2003US-0476189P.
XX (ARBO-) ARBorgen LLC.
XX Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX
DR WPI; 2005-075542/08.
DR P-PSDB; ADW17479.
XX
XX New polynucleotides isolated from plants encoding transcription factors,
PT and polypeptides encoded by such polynucleotides, useful for regulating
PT gene transcription and gene expression.
XX
PS Claim 3; SEQ ID NO 421; 1265pp; English.
XX
XX This invention relates to novel isolated plant nucleic acid molecules, or
CC variants thereof, that encode transcription factors. Specifically, it
CC refers to transcription factor proteins that are capable of binding to
CC DNA in order to regulate gene transcription and gene expression in a
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC invention describes DNA constructs containing DNA encoding a
CC transcription factor that regulates the promoter, which is operably
CC linked to the desired nucleic acid to be expressed. It further provides
CC transgenic plants expressing a transcription factor that confers a trait
CC to the plant such as increased drought, salt or disease tolerance, height
CC change, enhanced cold/frost tolerance, enhanced color, health and
CC nutritional characteristics, as well as improved taste, starch
CC composition, flower longevity and germination, amongst others.
CC Accordingly, such plants that are successfully transfected with a DNA
CC construct can be characterized by a difference in flower color, petal or
CC leaf shape and size, aroma or plant height. This polynucleotide is a
CC plant transcription factor cDNA sequence of the invention.
XX
SQ Sequence 1681 BP; 381 A; 418 C; 464 G; 418 T; 0 U; 0 Other;

Query Match 34.0%; Score 387.4; DB 14; Length 1681;
Best Local Similarity 62.9%; Pred. No. 2.7e-116;
Matches 690; Conservative 0; Mismatches 386; Indels 21; Gaps 5;

QY 61 GCGGCTCAACCGCTAATAGTTTCAGACCTTTTCAGCTGGTCCGAGCTCCCGAGGAGCTG 120
DB 285 GCGGCGCGGAGCGAGACGGCTTGAGCCGAGCTGCTGGGCCCAACATGCGCAGAGCTG 344

QY 121 CTTAGAGAAATCCTGATTAGGGTTGAGACTGTGTTGACGGCGGCGATGTCGCGCGCA 180
DB 345 CTGCGGAGGTTCTGTTGAAGATCGAGGCGTCCGA---GGGGTCTGTCGCGCGGAAG 401

822 TAACTCTTGGGACCAATTTACTATCTATGATGCCATCCACCTTATGCTGGAGATCT 881
566 -----CGAAGATGCGAAGAGCGGCTCTCTTAATTTTCATCAAAAGTTTACCTAGAGTTCC 620
882 GATTTTCAAAAGGGTCAGTCTGCGACGTGTGATTTGGTTCAAAACCATTTTATCCCCAGGATACC 941
621 TCAGGGAAGTTACCCCATCGCTCACATTTTCATACGAGTTTAAACGTTCTTAGGCTCTCGGG 680
942 TCGTGGGAATATCCAGTCTCAATATTTTCATAGAAATTTGAACGTTTGGGTTCTAGGG 1001
681 ACCGAGAAGATGCGTTGTCATCATGATGATACAATACCTATGAGCATCGTGGAGTCGCGAGG 740
1002 CCGAGAGAGATGCACTGCGCTATGATTTCTATCCCTGTATCAGCAATTTGACAGGAGG 1061
741 AGTAGTAGTTTCAACATCATAGCTCTTTTCCAGTCGGTATCATACCAAGTCTTTAGGTC 800
1062 AACAGCTCCGACGACACTGAAATTTCTTTGAGTTACCATGAATCTTTTCCATCAATTC 1121
801 TCACTCAAAACCAATTCGCGAGTAATAGTGCATCATGTAGGACTCAGGCAACACCTGGG 860
1122 CTTTTTCAAAATCAAAATCAGTTGAGCGAATAATTCAACTGCTTCAATTACTCACTCAAAA 1181
861 AGATCCACCACTGGTCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTACGTTGCTG 920
1182 TGAAGCAAACTGGTGTGAGNACAAAGTCCCGAGGTGGCAGACATCTGCGAGTCTG 1241
921 GTGCTTAAATTTCCATGGTGCAGTGCACAGTGGCTTCGGTTAAGAACTTTACGCTGTGGC 980
1242 GTGCTTAACTTCCATGGAGGGTAAAGCTTGCATCGGTAAAGAACTTCCAGCTGGTGGC 1301
981 AGTTAGTGACTGTGA---AGCAGGGCAGACATCTGAGAGGATCATACTCCAGTTTGGGAA 1037
1302 TTCAGATGAGAGTAACCCCAACCAACGAGCAACGATGATGTGATCTCCAGTTTGGTAA 1361
1038 AGTTGGGAAGGACATGTTTACCAGTATGATGATATCGATTTCTGCGTTTCAACGCTT 1097
1362 AGTTGGGAAGGACATGTTTACCAGTATGATGATATCGGCTATCTTCTGCGCAATTTGAGCGTT 1421
1098 TGCTATCTGCTGAGCAGTTTGAACACCAAGAAATTCGCTGTGAA 1140
1422 TGGCATATGCTGAGCAGCTTTGACACCAAAATAGGTGGAA 1464

RESULT 13
ID ABL93451/C
XX ABL93451 standard; cDNA; 469 BP.
AC ABL93451;
XX
XX 10-JUN-2002 (first entry)
DT
XX Arabidopsis thaliana nucleic acid sequence Ref:2027216 SEQ ID NO:216.
DE
XX Arabidopsis thaliana; insecticide; fungicide; plant; mapping; diagnosis;
KW Arabidopsis thaliana; insecticide; gene; ss.
XX genetic modification; gene; ss.
XX Arabidopsis thaliana.
XX
XX US2002023280-A1.
XX
XX 21-FEB-2002.
PD
XX
XX 26-JAN-2001; 2001US-00770444.
PF
XX
XX 27-JAN-2000; 2000US-0178502P.
PR
XX
XX (GORL/) GORLACH J.
PA (ANY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.

(RAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX WPI; 2002-267486/31.
DR
XX
XX New Arabidopsis thaliana nucleic acid, for identifying homologous genes,
PT producing compositions that modulate the expression or function of its
PT encoded protein, and mapping functional regions of a protein.
XX
PS Claim 1; SEQ ID NO 216; 44pp; English.
XX
CC The present invention describes an Arabidopsis thaliana nucleic acid (1)
CC comprising a sequence capable of hybridising under stringent conditions
CC to a sequence (S1) selected from any one of the 999 sequences given in
CC ABL93236 to ABL94234. (1) have insecticide and fungicide activities, and
CC they can be used as protein expression modulators. (1) can be used in
CC identifying homologous or related genes, in producing compositions that
CC modulate the expression or function of their encoded proteins, mapping
CC functional regions of the proteins, and in studying associated
CC physiological pathways. (1) can also be used: (1) for the genetic
CC manipulation of cells, particularly plant cells; (2) in screening assays
CC of various plant strains to determine the strains that are best capable
CC of withstanding a particular disease or environmental stress; (3) for
CC enhancing or inhibiting production of a biosynthetic product in a plant;
CC (4) as probes in mapping and in diagnosis, in genetic modification and
CC for screening purposes, to generate additional copies of the nucleic
CC acids, to generate ribozymes or antisense oligonucleotides, and as single
CC stranded DNA probes or as triple-strand forming oligonucleotides; and (5)
CC for generating genetically modified transgenic organisms. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site
XX
SQ Sequence 469 BP; 133 A; 116 C; 90 G; 130 T; 0 U; 0 Other;
Query Match 28.6%; Score 326; DB 6; Length 469;
Best Local Similarity 100.0%; Pred. No. 2.4e-96;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 815 TCGCAGTAATAGTGCATCATCTAGGACTCAGGCAACCACTGGAGATCCACCATGG 874
DB 469 TCGCAGTAATAGTGCATCATCTAGGACTCAGGCAACCACTGGAGATCCACCATGG 410
QY 875 TGCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTAGCTTGCTGGTCTTAAATTTCC 934
DB 409 TGCTGAGCAACAAAGCTCCACGGTGGCATGAGCAGTTAGCTTGCTGGTCTTAAATTTCC 350
QY 935 ATGGTTCGAGTCAAGTGGCTTCGGTTAAGAACTTTTCAGCTTGTGGCAGTTAGTACTGTG 994
DB 349 ATGGTTCGAGTCAAGTGGCTTCGGTTAAGAACTTTTCAGCTTGTGGCAGTTAGTACTGTG 290
QY 995 AAGCAGGGCAGACATCTGAGAGATCATACTCCAGTTTGGGAAGTTGGGAAGGACATGT 1054
DB 289 AAGCAGGGCAGACATCTGAGAGATCATACTCCAGTTTGGGAAGTTGGGAAGGACATGT 230
QY 1055 TTACCATGATTTATGGATATCCGATTTCTGCGTTTCAACGCTTTGCTATCTGCTCAGCA 1114

RESULT 15
ADW17066
ID ADW17066 standard; cDNA; 1923 BP.
XX AC
XX ADW17066;
XX
XX 24-MAR-2005 (first entry)
XX DE
XX DE Pinus radiata transcription factor cDNA TUBBY family Seq 809.
XX KW
XX KW gene; ss; plant; transcription; gene regulation; gene expression;
XX KW transgenic plant; drought resistance; disease resistance; salt tolerance;
XX KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
XX KW flower color.
XX
XX OS
XX OS Pinus radiata.
XX PN
XX PN WO2005001050-A2.
XX
XX PD
XX PD 06-JAN-2005.
XX PF
XX PF 07-JUN-2004; 2004WO-US017965.
XX PR
XX PR 06-JUN-2003; 2003US-0476189P.
XX PA
XX PA (ARBO-) ARBorgen LLC.
XX PI
XX PI Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
XX PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
XX PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX
XX WPI; 2005-075542/08.
XX DR
XX DR P-PSDB; ADW17851.
XX
XX PT
XX PT New polynucleotides isolated from plants encoding transcription factors,
XX PT and polypeptides encoded by such polynucleotides, useful for regulating
XX PT gene transcription and gene expression.
XX
XX PS
XX PS Claim 3; SEQ ID NO 809; 1265pp; English.
XX
XX CC
XX CC This invention relates to novel isolated plant nucleic acid molecules, or
XX CC variants thereof, that encode transcription factors. Specifically, it
XX CC refers to transcription factor proteins that are capable of binding to
XX CC DNA in order to regulate gene transcription and gene expression in a
XX CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
XX CC invention describes DNA constructs containing DNA encoding a
XX CC transcription factor that regulates the promoter, which is operably
XX CC linked to the desired nucleic acid to be expressed. It further provides
XX CC transgenic plants expressing a transcription factor that confers a trait
XX CC to the plant such as increased drought, salt or disease tolerance, height
XX CC change, enhanced cold/ frost tolerance, enhanced color, health and
XX CC nutritional characteristics, as well as improved taste, starch
XX CC composition, flower longevity and germination, amongst others.
XX CC Accordingly, such plants that are successfully transfected with a DNA
XX CC construct can be characterized by a difference in flower color, petal or
XX CC leaf shape and size, aroma or plant height. This polynucleotide is a
XX CC plant transcription factor cDNA sequence of the invention.
XX
XX SQ
XX SQ Sequence 1923 BP; 514 A; 374 C; 457 G; 578 T; 0 U; 0 Other;

Query Match 28.2%; Score 322; DB 14; Length 1923;
Best Local Similarity 59.4%; Pred. No. 1.1e-94;
Matches 641; Conservative 0; Mismatches 400; Indels 39; Gaps 4;

QY 97 TGGTCGGAGCTCCCGGAGGAGCTGCTTAGAGAAATCCTGATTAGGCTTCGAGACTGTTGAC 156
DB 545 TGGGCGAATATGCCCTCGAGTTGATAGAGATGTTATTTCAGAGAAATTGAG---GAGAGC 601

QY 157 GCGCGGATTTGGCGGTTCGCGGCGAAACGTTGGTGGCTTGTCCCGCGCTTTGTCTAGCTGG 216
DB 602 GAGAGCAGTTGGCTTTTCAGGAGAGAGTGTGGTGGCATGTGCTGGGGTTGTATGACATGG 661

QY 217 AGGATTCTCACAAAGGAGATTGTAGCTGTCTCTGAATTTCTCTCTAAATTTGACTTTCCCT 276
DB 662 AGGCAAAATTACAAAGGAGCTTGTCCGATACCTGAGCTATCTGGGAAGCTTACCTTCCCC 721

QY 277 ATCTCCCTCAAGCAGTCTGGTCCCAAGAGATTCTCTAGTTCATATGCTTTTATAAAAGTAAAT 336
DB 722 ATTTCTTTGAAGCAGCTGGATCAAGGGAGCCCTCTGTACAGTGTCTTATCAAGCGAGAT 781

QY 337 CGAATACTCAATCGTATCATCTCTATCTCCGATTAACCTACCTCTTTGACGATCAACGG 396
DB 782 AGGGCTACATCTACCTATCAATTTGTTCTCTGTCTGAATATCAACACTTTCTGAAATGA 841

QY 397 AAGTTTCTTTCTGTCTTCTTAAGCTGAAGCGCGCAACTTGCATCTGATTACATCTCTCT 456
DB 842 AAATTTCTTTCTTGGGCTCGGAATTCAGCGGGCTACAAGTACAGAGTATATTATTTCT 901

QY 457 TTGGCTTCAGACGATATCTCAAAAGAGAAGCAACGGTATCTTTGGGAGAGATGAGATCGAAC 516
DB 902 CTTAATGCAGACGATATGTCAAGGGGTAGCAATACATATGTTGGAAATTTGAGGTCAAC 961

QY 517 TTCCTTGGAAACAAAATTCACGGTCTTTGATGG-----TAGTCAGACCCGAGCA 564
DB 962 TTTCTTGGTACTAAGTTTACAATATATATAGTACCACCTCTCAAGTGGGCAATTTGCA 1021

QY 565 GCGAAGATGCAGAGAGCGGCTCTTCTAAATTTCAATCAAAAGTTTCACTAGAGTTCTCTCAG 624
DB 1022 TCACAAAGCGTGCAGGTGCGGGTAGGAGCAAAACAAGTTTCTTCCAAGAGTTCACGCT 1081

QY 625 GGAAGTTACCCCATCGCTCACATTTTCATACGAGTTAAAGCTTTAGGCTCTCGGGACCG 684
DB 1082 GGTAGCTATAATATTGCTCACATTTGGCTATGAATTTGAATGTCTTGGGAACTAGGGGTCCA 1141

QY 685 AGAAGAAATCGTTGCATCATGATGATACATCTATGAGCATGCTGGAGTCGCGAGGAGTA 744
DB 1142 AGACGATGCAATGATGATCATGATCTTATTCAGCTTCTGCTGTGAGCAGAGGAGCAGT 1201

QY 745 GTAGCTTCAACATCCATAAGCTCTTTTCCAGTCCG-----TCATCACCACCA 789
DB 1202 GCCCAACTCTCTGTTGAACCTTCGCTCGTGTAGTCTAGACCAGTCTCTTGCACTTTGCTCT 1261

QY 790 GTCTTTAGGTCTCACTCAAAACCATTTGCGCAGTAAATAGTGCATCATGTAGCGACTCAGGC 849
DB 1262 ACTTTTGGCTCCAAAGTCTGCTGCAATTTTAGAAAGTTTCATCTCTGAGGCCACTTACCGGT 1321

QY 850 AACACCTGGGAGATCCACCATTTGCTGAGCAACAAGCTCCACGCTGGCATGAGCAG 909
DB 1322 AGCAACTTGAAGATGAACCATTTGGTTTAAAAAAATAAGCGCCAAAGTGGCATGAGCAG 1381

QY 910 TTACGCTGCTGCTGCTTAAATTTTCCATGCTGAGTCACAGTGGCTTCGGTTAAGAACTTT 969
DB 1382 CTTCAATGTTGGTGTGTTGAATTTCAAGGGGCTGTGACTGTGGCTTCTGTTAAGAAATTT 1441

QY 970 CAGCTTGTGGCAGTTAGTGAATGTGAAGCA-----GGCGAGACATCTGAGAGGATC 1020
DB 1442 CAGCTTGTGTCAGCACCTTGAGCCTAATCCATCGGTTTCTCAGAATGATTGTGATAAGGTC 1501

QY 1021 ATACTCCAGTTTGGGAAAGTTGGGAAGACATGTTTACCATGATTTATCGATTCGATTCGATT 1080
DB 1502 ATACTCCAATTTGGAAAGGTTGGAAAGGACATGTTTCACAATGATTTATCGTACCCGCTC 1561

QY 1081 TCTGGGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTTGAACCCAGAAATTCGCTGTGAA 1140
DB 1562 TCAGCTTTTCAAGCTTTTGGCCATATGCTTAGTAGTTTGTGACACAAACTTGTCTGTGAA 1621

Search completed: December 26, 2005, 03:41:22
Job time : 803 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 01:47:21 ; Search time 6049 Seconds

(without alignments)
10712.768 Million cell updates/sec

Title: US-10-763-042-20

Perfect score: 1140

Sequence: 1 atgagcttcgaagttact.....aaaccagaattgcctgtgaa 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.cm.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sv.*
12: gb.un.*
13: gb.vi.*
14: gb.htg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1140	100.0	1143	6	CQ804532 Sequence
2	1140	100.0	1143	15	AF487270 Arabidops
3	1037.4	91.0	1595	15	BT004092 Arabidops
C	598	52.5	86022	15	ATAC011623 Arabidops
	572.8	50.2	1170	15	BT020601 Arabidops
6	572.8	50.2	1259	15	BT010768 Arabidops
7	567.8	49.8	1143	15	AY046922 Arabidops
8	350.6	30.8	1781	15	BT013601 Lycopersi
9	342.6	30.1	1221	6	CQ804888 Sequence
10	342.6	30.1	1221	15	AY045774 Arabidops
11	342.6	30.1	1252	15	AY096504 Arabidops
12	342.6	30.1	1943	15	AY074273 Arabidops
13	336.2	29.5	1687	15	AK071159 Oryza sat
14	336.2	29.5	1948	15	AK060587 Oryza sat
15	262.8	23.1	2102	15	AK102370 Oryza sat
16	250.4	22.0	1185	15	AY045773 Arabidops
17	248.8	21.8	1699	15	BT008627 Arabidops
18	240.4	21.1	1571	15	AK062089 Oryza sat

19	240.4	21.1	2345	15	AK064855	Oryza sat
20	210.2	18.4	1167	15	AF487268	Arabidops
21	210.2	17.6	1722	15	CAR400860	Cicer ari
22	195.4	17.1	1338	15	AF487271	Arabidops
23	193	16.9	1140	15	AY092403	Arabidops
24	193	16.9	1171	15	AY059088	Arabidops
25	193	16.9	1388	15	AF370146	Arabidops
26	187.4	16.4	85746	15	AC051627	Genomic S
27	184.2	16.2	878	15	AF386521	Pyrus com
28	184	16.1	1290	15	AY046921	Arabidops
29	184	16.1	1321	15	AY122972	Arabidops
30	184	16.1	1567	15	AF424565	Arabidops
31	184	16.1	1582	15	AY072154	Arabidops
32	179.8	15.8	1539	15	AB023895	Lemna pau
33	176	15.4	1341	15	AK102298	Oryza sat
34	171.2	15.0	2027	15	AK103583	Oryza sat
35	170.2	14.9	1103	15	BT018325	Zea mays
36	169.6	14.9	2023	15	AK069622	Oryza sat
37	169.2	14.8	1257	15	AK106164	Oryza sat
38	165	14.5	1630	15	AK098851	Oryza sat
39	165	14.5	1685	15	AK061747	Oryza sat
40	163.4	14.3	2129	15	AK100078	Oryza sat
41	163.4	14.3	3119	15	AK120706	Oryza sat
42	160.8	14.1	1824	15	AK066910	Oryza sat
43	160.4	14.1	1994	15	AK102221	Oryza sat
44	159.2	14.0	1800	15	AK104333	Oryza sat
45	159.2	14.0	1826	15	AK098928	Oryza sat

ALIGNMENTS

RESULT 1
CQ804532
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Sequence 943 from Patent WO2004035798.
CQ804532
CQ804532.1 GI:47110900
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

Inze, D., de Veylder, L. and Vlieghe, K.

Identification of novel e2f target genes and use thereof

Patent: WO 2004035798-A 943 29-APR-2004;

CropDesign N.V. (BE)

Location/Qualifiers

source
1. .1143
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match	100.0%;	Score 1140;	DB 6;	Length 1143;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1140;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGACGTTCCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTACGCC	60	
Db	1	ATGACGTTCCGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTACGCC	60	
Qy	61	GCGCCTCAACCGCTAATAGTTTTCAGCTGTCGAGCTCCGCGAGAGGAGCTG	120	
Db	61	GCGCCTCAACCGCTAATAGTTTTCAGCTGTCGAGCTCCGCGAGAGGAGCTG	120	
Qy	121	CTTAGAGAAATCCTGATTAGGGTTGAGACTGTTGACGCGCGGCGATTGGCGCTCGCGCGA	180	
Db	121	CTTAGAGAAATCCTGATTAGGGTTGAGACTGTTGACGCGCGGCGATTGGCGCTCGCGCGA	180	
Qy	181	AACGTGGTGGCTTGTGCGCGGCTTTTGTGTAGTGGAGGATTTCTACCAAGGAGATTGTA	240	

[illegible]

QY	421	CTGAAGCGCGCAACTTGGCACTGATTCATCATCTCTTTGGGTTTCAGACGATATCTCAAG	480
Db	421	CTGAAGCGCGCAACTTGGCACTGATTCATCATCTCTTTGGGTTTCAGACGATATCTCAAG	480
QY	481	AGAAAGCAACGCGTATCTTGGGAGATGAGATCGAATCTCTTGGAAACAAATTCACGGTC	540
Db	481	AGAAAGCAACGCGTATCTTGGGAGATGAGATCGAATCTCTTGGAAACAAATTCACGGTC	540
QY	541	TTTGATGTAGTCAGACCGGAGCGAAGATGAGAGAGCGCGCTCTTCTTAATTTTCATC	600
Db	541	TTTGATGTAGTCAGACCGGAGCGAAGATGAGAGAGCGCGCTCTTCTTAATTTTCATC	600
QY	601	AAAGTTTCACTTAGAGTTCTCAGGAGTACCCCATCGCTCACATTTTCATACGATTA	660
Db	601	AAAGTTTCACTTAGAGTTCTCAGGAGTACCCCATCGCTCACATTTTCATACGATTA	660
QY	661	AAAGTCTTAGGCTCTCGGAGCGGAGAAAGTGGTGGATCATCATGATCAATACCTATG	720
Db	661	AAAGTCTTAGGCTCTCGGAGCGGAGAAAGTGGTGGATCATCATGATCAATACCTATG	720
QY	721	AGCATCGTGGAGTCCGAGGAGTAGCTTCAACATCCATAGCTCTTTTCCAGTCGG	780
Db	721	AGCATCGTGGAGTCCGAGGAGTAGCTTCAACATCCATAGCTCTTTTCCAGTCGG	780
QY	781	TCATCACAGTCTTTAGTCTCACTCAAAACCATTCGCGAGTAATAGTCATCATGTAGC	840
Db	781	TCATCACAGTCTTTAGTCTCACTCAAAACCATTCGCGAGTAATAGTCATCATGTAGC	840
QY	841	GACTCAGGCAACCACTCGGAGATCCACCATTTGCTGAGCAACAAAGCTCCACGGTCG	900
Db	841	GACTCAGGCAACCACTCGGAGATCCACCATTTGCTGAGCAACAAAGCTCCACGGTCG	900
QY	901	CATGAGCAGTTAGCTTGGTCTTAAATTTCCATGTCAGTCCAGTGGCTTCGGTT	960
Db	901	CATGAGCAGTTAGCTTGGTCTTAAATTTCCATGTCAGTCCAGTGGCTTCGGTT	960
QY	961	AGAACTTTTCAAGCGTTTGTAGTGTGTAAGAGCGAGGAGAGTCTGAGGAGATC	1020
Db	961	AGAACTTTTCAAGCGTTTGTAGTGTGTAAGAGCGAGGAGAGTCTGAGGAGATC	1020
QY	1021	ATACTCCAGTTTGGGAAAGTGGGAGGACATGTTTACCATGATATGATATCCGATT	1080
Db	1021	ATACTCCAGTTTGGGAAAGTGGGAGGACATGTTTACCATGATATGATATCCGATT	1080
QY	1081	CTGCGTTTCAAGCGTTTGTATCTGCTGAGCAGTTTGAACACCAAGATTGCTGTGAA	1140
Db	1081	CTGCGTTTCAAGCGTTTGTATCTGCTGAGCAGTTTGAACACCAAGATTGCTGTGAA	1140
RESULT 3			
BT004092			
LOCUS	BT004092 1595 bp mRNA linear PLN 14-FEB-2003		
DEFINITION	Arabidopsis thaliana clone RAFL16-03-H24 (R20922) At3g06380 mRNA, complete sequence.		
ACCESSION	BT004092		
VERSION	BT004092.1 GI:28393389		
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 1595)		
AUTHORS	Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Arabidopsis Full Length cDNA Clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1595)		
AUTHORS	Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-FEB-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA		
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.		
FEATURES	The Salk, Stanford, PGENC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.		
source	Yamada,K. (SSP/PGENC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGENC) contributed equally to this work as PIs.		
Location/Qualifiers	Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.		
1. .1595			
/organism="Arabidopsis thaliana"			
/mol_type="mRNA"			
/db_xref="taxon:3702"			
/chromosome="3"			
/clone="RAFL16-03-H24 (R20922)"			
/ecotype="Columbia"			
/note="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert."			
gene	1. .1595		
	/gene="At3g06380"		
misc_feature	147. .1380		
	/gene="At3g06380"		
	/note="This is a potential full length cDNA corresponding to At3g06380 and may be the product of a partially processed transcript."		
	/evidence=experimental		
misc_difference	1016		
	/gene="At3g06380"		
	/note="compared to genomic sequence"		
	/replace="g"		
ORIGIN			
Query Match	91.0%; Score 1037.4; DB 15; Length 1595;		
Best Local Similarity	92.5%; Pred. No. 0;		
Matches 1139; Conservative	0; Mismatches 1; Indels 91; Gaps 1;		
QY	1	ATGACGTTCCGAAGTTTACTCCAGGAAATGCGGTCTTAGCCACACCGTGTAGTTACGCC	60
Db	147	ATGACGTTCCGAAGTTTACTCCAGGAAATGCGGTCTTAGCCACACCGTGTAGTTACGCC	206
QY	61	GCGGCTCAACCGCTAAATAGTTTCAGACCTTTTCAGTGGTCCGAGGAGGAGCTG	120
Db	207	GCGGCTCAACCGCTAAATAGTTTCAGACCTTTTCAGTGGTCCGAGGAGGAGCTG	266
QY	121	CTTAGAGAAATCCTGATTAGGTTGAGACTGTGACGCGCGGAGATGGCGTCCGCGGA	180
Db	267	CTTAGAGAAATCCTGATTAGGTTGAGACTGTGACGCGCGGAGATGGCGTCCGCGGA	326
QY	181	ACGTGGTGGCTTGGCGCGGCTTTGCTGCTAGCTGAGGATTTCTACCAAGGAGATTGA	240

Db 327 AACGTGGTGGCTTGTGCGCGCGTTTGTCTAGCTGGAGGATTTCTCACCAGGAGATTGTA 386
Qy 241 GCTGTTCTCGAATTCCTCTAAATGACTTTCCTATCTCCCTATCTCCCTCAAGCAGTCTGGTCCA 300
Db 387 GCTGTTCTCGAATTCCTCTAAATGACTTTCCTATCTCCCTATCTCCCTCAAGCAGTCTGGTCCA 446
Qy 301 AGAGATTCTCTAGTTCAATAGCTTTATAAACGTAATCGAAATACCTCAATTCGTATCATCTC 360
Db 447 AGAGATTCTCTAGTTCAATAGCTTTATAAACGTAATCGAAATACCTCAATTCGTATCATCTC 506
Qy 361 TATCTCGGATTAATACCTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTTCTTAAG 420
Db 507 TATCTCGGATTAATACCTCTTTGACGGATAACGGGAAGTTTCTTCTGCTGCTTCTTAAG 566
Qy 421 CTGAGCGCGCAACTGCACTGATTAACATCATCTCTTTGCGGTTCAGCAGATATCTCAAG 480
Db 567 CTGAGCGCGCAACTGCACTGATTAACATCATCTCTTTGCGGTTCAGCAGATATCTCAAG 626
Qy 481 AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTGGAAACAAAATTCACGGTTC 540
Db 627 AGAAGCAACGCGTATCTTGGGAGATGAGATCGAACTTCTTGGAAACAAAATTCACGGTTC 686
Qy 541 TTTGATGTTAGTCAGACCGGAGCAGCAAGATGCAAGAGCGGCTCTTTCTAATTTTCATC 600
Db 687 TTTGATGTTAGTCAGACCGGAGCAGCAAGATGCAAGAGCGGCTCTTTCTAATTTTCATC 746
Qy 601 AAAGTTTACCTAGATTCTCTAGGGAAGTTACCCATCGCTCACATTTTCATCAGATTTA 660
Db 747 AAAGTTTACCTAGATTCTCTAGGGAAGTTACCCATCGCTCACATTTTCATCAGATTTA 806
Qy 661 AAGCTCTTAGGCTCTC----- 676
Db 807 AAGCTCTTAGGCTCTCGTTAAACAGTTTTCATTCAGTATGATTAATGCTTCTTCTAGT 866
Qy 677 -----GGGACCGGAGAG 689
Db 867 CTTTGGACCATCTGATTTTAAACAGTCTTTTGAGGGTATTAAATAACAGGGGACCGAGAAG 926
Qy 690 AATGCGTTGCATCATGATTAACATCTATGAGCATCTGTTGAGTCCGAGGAGTAGTAGC 749
Db 927 AATGCGTTGCATCATGATTAACATCTATGAGCATCTGTTGAGTCCGAGGAGTAGTAGC 986
Qy 750 TTCACATCATTAAGCTCTTTTCCAGTCCGTCATCACCAAGTCTTTTGTCTTCACTCAAA 809
Db 987 TTCACATCATTAAGCTCTTTTCCAGTCCGTCATCACCAAGTCTTTTGTCTTCACTCAAA 1046
Qy 810 ACCATTGCGCAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACTGGGAGATCCACC 869
Db 1047 ACCATTGCGCAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACTGGGAGATCCACC 1106
Qy 870 ATTGTTGCTGAGCAACAAAGCTCCAGGTGGCATGAGCATTAAGTTGCTGTTGCTTAA 929
Db 1107 ATTGTTGCTGAGCAACAAAGCTCCAGGTGGCATGAGCATTAAGTTGCTGTTGCTTAA 1166
Qy 930 TTTCCATGTTGAGTCCAGTCCAGTGGCTTTCGGTTAAGAACTTTTCAGTTTGTGCGAGTTAGTGA 989
Db 1167 TTTCCATGTTGAGTCCAGTGGCTTTCGGTTAAGAACTTTTCAGTTTGTGCGAGTTAGTGA 1226
Qy 990 CTGTGAAGCAGGGCAGACATCTGAGAGGATCATCTCAGATTGGGAAAGTTGGGAAGGA 1049
Db 1227 CTGTGAAGCAGGGCAGACATCTGAGAGGATCATCTCAGATTGGGAAAGTTGGGAAGGA 1286
Qy 1050 CATGTTTACATGGAATATGATATCCGATTTCTGGGTTTCAAGCGTTTGTCTATCTGCCT 1109
Db 1287 CATGTTTACATGGAATATGATATCCGATTTCTGGGTTTCAAGCGTTTGTCTATCTGCCT 1346
Qy 1110 GAGCAGTTTGAACACAGATTGCTGTGAA 1140
Db 1347 GAGCAGTTTGAACACAGATTGCTGTGAA 1377

RESULT 4

ATAC011623/c

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

ATAC011623 86022 bp DNA linear PLN 30-OCT-2002
Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence,
complete sequence.
AC011623
AC011623.5 GI:12408729
HTG.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 86022)
Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.,
Rowman,C.L., White,O., Fujii,C.Y., Uterback,T.R., Barnstead,M.E.,
Bowman,C.M., Koo,H., Nierman,W.C. and Fraser,C.M.
Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence
Unpublished
2 (bases 1 to 86022)
Lin,X. and Kaul,S.
Direct Submission
Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org
3 (bases 1 to 86022)
Lin,X.
Direct Submission
Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Jan 24, 2001 this sequence version replaced gi:12280803.
Address all correspondence to:
Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr.
Rockville, MD 20850, USA
e-mail: xlin@tigr.org
BAC clone F24P17 is from Arabidopsis chromosome III and is near the
molecular marker mi403.
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.

Genes were identified by a combination of three methods: Gene
prediction programs including GRL (available by anonymous ftp
from arthur.epm.ornl.gov), Genefinder (Phil Green, University of
Washington), Genscan (Chris Burge,
http://www.cba.stanford.edu/~chris/GENSCANW.html), and NetPlantGene
(http://www.cba.stanford.edu/~chris/genetpgene.html), searches of the
complete sequence against a peptide database and the Arabidopsis
EST database at IIGR (http://www.tigr.org/tdb/at/at.html).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as 'unknown' proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as 'hypothetical' proteins. Genes encoding tRNAs are
predicted as tRNAcan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAcan-SE/). Simple repeats are
identified by repeatmasker (Arian Smit,
http://genome.washington.edu/BW/RepeatMasker.html). Regions of
genomic sequence that are not annotated as genes but have predicted
exons by GRL are annotated as misc features.

FEATURES
source

1. .86022
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="III"
/map="mi403"
/clone="F24P17"
/ecotype="Columbia"
-283. 1822
/gene="F24P17.1"
/note="similar to pyruvate dehydrogenase kinase
GB:AAC97601, [Arabidopsis thaliana]"
Join(<283. .529,638. .818,916. .1063,1101. .1142,1160. .1376,

gene

mRNA

Db 55732 AACGTGGTGGCTTGTGCGCGGCTTGTGCTAGCTGGAGGATTTCTCACCAGAGAGATTGTA 55673
QY 241 GCTGTTCCCTGAATTCCTCTAAATTTGACTTTCCCTATCTCCCTCAA----- 287
Db 55672 GCTGTTCCCTGAATTCCTCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGGTTCTCTGT 55613
QY 288 ----- 287
Db 55612 AAATTTTGATTTTGGAGTATAAGTGATCGAATTGAACAATAAAGTTAGTCTCCTTTACTT 55553
QY 288 ----- 287
Db 55552 CATCTCATAACATATGACAATTGAGTAGCGTTTGGCGTGGCTTTTAAAGTGTCTTGTGTC 55493
QY 288 ----- 287
Db 55492 TCAGCTCAACAAAAGCTTTTAGTCTCTTTGATCATTTGATATATAGATAAAGTTTCTTG 55433
QY 288 -----GCAGTCTGGTCCCAAGAGA 305
Db 55432 GTTTTGTAAAGCTGAAGTATTGGTACTTTTGAATGATATATATGAGTCTGGTCCCAAGAGA 55373
QY 306 TTCTCTAGTTCAATGCTTTTATAAAGCTAATCGAATACTCAATGATCATCTCTATCT 365
Db 55372 TTCTCTAGTTCAATGCTTTTATAAAGCTAATCGAATACTCAATGATCATCTCTATCT 55313
QY 366 CGGATTAACACTCT----- 379
Db 55312 CGGATTAACACTCTCTGTTCTCTCTAGATTTATCAAGAAGAGATTA 55253
QY 380 -----CCT 382
Db 55252 CTTGCTTTCTGAGTTCTGAAATGTTAAACAAGAGTCTCTGTCTAATTTTCTGACAGCTT 55193
QY 383 TGACGAGTAAACGGGAAGTTTCTTTCTGCTCTTCTAAGCTGGAAGCGCAACTTGCACGT 442
Db 55192 TGACGAGTAAACGGGAAGTTTCTTTCTGCTCTTCTAAGCTGGAAGCGCAACTTGCACGT 55133
QY 443 ATTACATCATCTCTTTGCGTTTCAGACGATATCTCAAAGAGAACACGGCTATCTTTGGGA 502
Db 55132 ATTACATCATCTCTTTGCGTTTCAGACGATATCTCAAAGAGAACACGGCTATCTTTGGGA 55073
QY 503 GAATG----- 507
Db 55072 GAATGAGTTTGCATATAAAGTTATCTCTTTTCGATATATAAGTATGCGATTTTCCAAAT 55013
QY 508 -----AGATCGAACTTCTTTGGAAACAAAATTCACGGTCTTTGATGTT 549
Db 55012 TTCTTATTGTTATTGCAAGATCGAATCTCTTTGGAACAAAATTCACGGTCTTTGATGTT 54953
QY 550 AGTCAGACCGGAGCAGCGAAGATGCAAGAGCGGCTCTTTCTAATTTCAATCAAAAGTTTCA 609
Db 54952 AGTCAGACCGGAGCAGCGAAGATGCAAGAGCGGCTCTTTCTAATTTCAATCAAAAGTTTCA 54893
QY 610 CTTAGAGTTCTCAGGGAAGTATCCCATCGCTCACAATTTCAACGAGTTTAAACGCTTAA 669
Db 54892 CTTAGAGTTCTCAGGGAAGTATCCCATCGCTCACAATTTCAACGAGTTTAAACGCTTAA 54833
QY 670 GGCTCTC----- 676
Db 54832 GGCTCTCGGTAAACAGCTTTTTCATTCAGTATGATAATGCTTCTTCTAGTCTTTTGACC 54773
QY 677 -----GGGGAACCGAAGAAATGCGTTG 698
Db 54772 ATTCTGATTTTAAACAGTCTTTTGAGGGTATTAATTAACAGGGGACCGAAGAAATGCGTTG 54713
QY 699 CATCATGATACATACCTATGAGCATCGTGGAGTCCGAGGAGTAGTACTTCAACATC 758
Db 54712 CATCATGATACATACCTATGAGCATCGTGGAGTCCGAGGAGTAGTACTTCAACATC 54653
QY 759 CATAAGCTCTTTTCCAGTCCGCTCATCACCAAGTCTTTTAGTGTCTCACTCAAAACCATTTGCG 818
Db 54652 CATAAGCTCTTTTCCAGTCCGCTCATCACCAAGTCTTTTAGTGTCTCACTCAAAACCATTTGCG 54593

QY 819 CAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACCATTTGGTGT 878
Db 54592 CAGTAATAGTGCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACCATTTGGTGT 54533
QY 879 GAGCAACAAAAGCTCCACGGTGGCATGAGCAGTTACGTTTGTGCTGCTTAAATTTCCATGG 938
Db 54532 GAGCAACAAAAGCTCCACGGTGGCATGAGCAGTTACGTTTGTGCTGCTTAAATTTCCATGG 54473
QY 939 TCGAGTCAAGTGGCTTTCGGTTAAGAACTTTTCAGCTTGTGGCAGTTAGTGAAGC 998
Db 54472 TCGAGTCAAGTGGCTTTCGGTTAAGAACTTTTCAGCTTGTGGCAGTTAGTGAAGC 54413
QY 999 AGGGCAGACATCTGAGAGGATCATACTCAGTTTGGGAAAGTTGGGAAGACATGTTTAC 1058
Db 54412 AGGGCAGACATCTGAGAGGATCATACTCAGTTTGGGAAAGTTGGGAAGACATGTTTAC 54353
QY 1059 CATGGATTATGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTT 1118
Db 54352 CATGGATTATGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGCTGAGCAGTTT 54293
QY 1119 TGAACACGAAATTCCTGTGAA 1140
Db 54292 TGAACACGAAATTCCTGTGAA 54271

RESULT 5
BT020601
LOCUS Arabidopsis thaliana 1170 bp mRNA linear PLN 05-FEB-2005
DEFINITION Arabidopsis thaliana At5g18680 gene, complete cds.
ACCESSION BT020601
VERSION BT020601.1 GI:58652097
KEYWORDS FLJ_CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE 1 (bases 1 to 1170)
AUTHORS Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.
TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1170)
AUTHORS Kim,C.J., Chen,H., Cheuk,R., Shinn,P. and Ecker,J.R.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2005) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
FEATURES Location/Qualifiers
1..1170
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U64136"
/ecotype="Columbia"
/note="This clone is in pUNI 51"
1..1170
/note="unknown protein"
/codon_start=1
/product="At5g18680"
/protein_id="AAW80874.1"
/db_xref="GI:58652098"
CDS
1..1170
/note="unknown protein"
/codon_start=1
/product="At5g18680"
/protein_id="AAW80874.1"
/db_xref="GI:58652098"
ORIGIN

Query Match		50.2%;	Score 572.8;	DB 15;	Length 1170;
Best Local Similarity		71.7%;	Pred. No. 2.5e-174;	Indels 69;	Gaps 5;
Matches 852;		Conservative 0;	Mismatches 267;		
QY	1	ATGACGCTTCGGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTCAAGCC	60		
DB	1	ATGACCTTAGCTAGCTTAATCTCTTGAGATGGTTCGAGACCGCATCGTGTGTCACGAC	60		
QY	61	-----GCCGCTCAACCGCTTAATAGTTTCAGACCCCTTTCAGCTGGTGC	102		
DB	61	CTTTCGCGCGCGCAGCTGCGGATTCCTCTGTGTCATCGCAAGATTATCGCTGGTCA	120		
QY	103	GAGCTCCCGAGGAGCTGCTTACAGAAATCTCATAGGTAGGTGAGACTGTTGACGGCGC	162		
DB	121	GAGATTCCTGAAGAGCTTCTTAGGGAGATTCTGATTCGTGTTGAACGGCGGACGCTGGC	180		
QY	163	GAATTGCGCGTGGCGGAAACGTGTGTGTGCGGGGTTTGTGCTAGCTAGGAGATT	222		
DB	181	GGATGGCGGTACAGCGCAGCGGTGTGTGTGCGGGGTTTGTGCTGGCTGGCGGCTA	240		
QY	223	CTCACCAAGAGATTGTAGCTGTCTCTGAATTCCTCTCTAAATGTATTCCTCTATCTCC	282		
DB	241	CTTATGAACGAAACCGTGTGTCTCTGAGATCTCTTCTAAAGTTGACTTTCGCCATCTCT	300		
QY	283	CTCAAGCAGTCTGCTCAAGAGATTCTCTAGTTCAATGCTTTTATAAAGTAAATCGAAT	342		
DB	301	CTCAAGCAGCTGCTCAAGAGATTCTCTAGTTCAATGCTTTTATAAAGTAAATCGAAT	360		
QY	343	ACTCAATCGTATCATCTCTATCTCGGATTAACCTCTCTTACGATTAACGGGAAGTTT	402		
DB	361	AGCATCATATCATCTCTATCTCGGATTAACCTCTCTTAAAGGATGATGGGAAGTTT	420		
QY	403	CTTCTTGTCTCTTAAAGTGAACCGCAACTTGTCACTGATTTACATCATCTCTTTGCGT	462		
DB	421	TTGCTTGTCTGCTGAAGTTGAAGCACCAACTTGTACGATTTACATTTATCTCTTACGT	480		
QY	463	TCAGACGATATCTCAAGAGAGACGCGGTATCTTGGGAGATGATGCAATTCCTT	522		
DB	481	TCTGATGATATGTGAGAGAGACCAAGCTTATGTTGGCAAGTGAATGCAATTCCTTA	540		
QY	523	GGAAACAAATTCACGCTCTTTGATGTGTAGTCAG-----ACCGGAGCAGCAAGATG	573		
DB	541	GGAACAAATTCACGCTCTTTGATGTGTAGTGAATCTGCTGCTTCAACGGAGCGCAAGTTG	600		
QY	574	CAGAAGCAGCGCTCTTCTAATTTTCATCAAGTTTACCTAGAGTTCTCTCAGGGAAGTTAC	633		
DB	601	AGAAAGACCGATCTTATAATCCGCAAGATTTCAGCAAAAGTTCTCTTGGAGTTAT	660		
QY	634	CCATCGCTCAATTTTCATACGAGTTAAACGTCTTAGGCTCTCGGGACCGGAGAAGTG	693		
DB	661	CCTGTCGCTCATATCATATGAGCTGAATGTCTTAGGATCCCGGGACCAAGAAAGATG	720		
QY	694	CGTTGCATCATGATACATACCTATGACATCGTGGAGTCGCGAGGTAGTGTAGCTTCA	753		
DB	721	CAATGTCTTATGACACAATACCTCAAGCAATGAGCCTCAAGAGTAGTCTCAGAA	780		
QY	754	ACATCATCAAGCTCTTTTCCAGTCCGTCATCAACAGTCTTTAGTCTCTCACTCAAAACCA	813		
DB	781	CCATCAGAGTTTCCCTTACTCGGTACTCTCGGTCAACCTTATCCAGGTCTCAGTCAAAACCA	840		
QY	814	TTGCGCAGTAATAGTGCATCATGTAGGACTCAGGCAACCACTGGGAGATCCACATTTG	873		
DB	841	TTACGCAGTA-----GCTCAAGCCACCTGAAGAAACACCATTA	879		
QY	874	GTGCTGAGCAACAAGCTCCAGGTGCGATGACGAGTTAGGTGCTGGTCTTAATTTTC	933		
DB	880	GTGCTGAGCAACAAGCACCAAGTGGCACGAGCAGCTACGCTGCTGGTCTTGAATTTTC	939		
QY	934	CATGTCGAGTACAGTGGCTTCGGTTTAAAGACTTTTCAGCTTGTGGCAG-----TT	984		
DB	940	CATGGCCGTGTACAGTACGCTCAGTGAAGAACTTTTCAGCTGTGGCAGCAGGAGTAGC	999		
QY	985	AGTGACTGTGAAGCAGGCGAGACATCTGAGA-----GGATCATACTCTCCAGTTT	1032		

Db	1000	TGTGGCAGTGGCACCGGAATGTACCGGAGAGCGAGCGAGCGAATATATTGCAAGTTT	1059	
Qy	1033	GGCAAAAGTTGGGAAGGACATGTTTACCATGGATTATGGATATCCGATTTCTGGCTTTCAA	1092	
Db	1060	GGCAAAAGTCCGGAAGATATGTTACGATGGATTATGGATACCGATCTCAGCTTTCCAG	1119	
Qy	1093	GCGTTTGCTATCTGCTGAGCAGTTTTGAAACAGAAATGCGCTGTGAA	1140	
Db	1120	GCTTTTGCCATTTGCTTGGACGACTTTGAGACTAGAAATCGCTTGTGAA	1167	
RESULT 6				
LOCUS	BT010768	Arabidopsis thaliana	1259 bp	mRNA linear
DEFINITION	BT010768	Arabidopsis thaliana At5g18680 gene, complete cds.		
ACCESSION	BT010768			
VERSION	BT010768.1	GI:38564317		
KEYWORDS	FLI CDNA.			
SOURCE	Arabidopsis thaliana	(thale cress)		
ORGANISM	Arabidopsis thaliana			
REFERENCE				
AUTHORS	Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.			
TITLE	Arabidopsis cDNA clones			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1259)			
AUTHORS	Cheuk,R., Chen,H., Kim,C.J., Shinn,P. and Ecker,J.R.			
TITLE	Direct Submission			
JOURNAL	Submitted (27-NOV-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA			
FEATURES				
source	1..1259 Location/Qualifiers			
	/organism="Arabidopsis thaliana"			
	/mol_type="mRNA"			
	/db_xref="taxon:3702"			
	/chromosomes="5"			
	/clone="Sf4136"			
	/ecotypes="Columbia"			
	/note="This clone is in pUNI 51"			
	32..1201			
	/notes="unknown protein"			
	/codon_start=1			
	/product="At5g18680"			
	/protein_id="AA023738.1"			
	/db_xref="GI:38564318"			
	/translation="MTLRSLILEMRSRPHRVVHDLAAAAADSTSVSSODYRWSEIPE ELLRDLTVQCFIKNNRTIQSYHLXILGTLNSLTDDGKFLLAACKLKH7TCTDYIILR OQPRDLISVQADGGWPSRVSACVAGWCRGWRLLMNETVVPVVISSKLTFPISLK SDMRSRSQAYGVKRSNPLGTFKTFVFDGNLLPSTGAALKRKSRYNPAKVSAPVIG SPKVAHVELNLVSGRPKKQCLMDTTPSTMEPQGVASPESEFPILGTRSLRS QSKPLRSSSHLKEPTVLNSLGRHEQLRCWLFNHRGTVASVKNFQLVNVAAGSAC GSGTGMSPERQSERIILQFGVKGMDFMTDYGYPISAFQAFACLSSFFETRIACE"			
CDS				
ORIGIN				
Query Match	50.2%; Score 572.8; DB 15; Length 1259;			
Best Local Similarity	71.7%; Pred. No. 2.5e-174;			
Matches 852; Conservative	0; Mismatches 267; Indels 69; Gaps 5;			
Qy	1	ATGACGTTCCGGAAGTTTACTCCAGGAAATGCGGTCTAGGCCACACCGTGTAGTTTCAAGCC	60	
Db	32	ATGACCTTAGCTAGCTTAATCTCTTGAGATGCGTTCGAGACCGCATCGTGTGTCACGAC	91	
Qy	61	-----GCCGCTCAACCGCTTAATAGTTTCAGACCCCTTTCAGCTGGTGC	102	
Db	92	CTTTCGCGCGCGCAGCTGCCGATTCACCTTCTGTGTCATCGCAAGATTATCGCTGTCA	151	
Qy	103	GAGCTCCCGGAGGCTGCTTAGAGAAATCTTGATAGGTTGAGACTGTTGACGGCGC	162	


```

Db      469 AGAAGAGCCAGCTTATGTTGCAAGTGGATCGAATCTCTAGGACGAATTCACCT 528
Qy      538 GTCTTTGATGGTAGTCAG-----ACCGAGCAGCGAAGATGCGAAGAGCCGCTCT 588
Db      529 GTCTTTGATGGAATCTGCTGCTTCAACCGGAGCGCCAAAGTTGAGAAAGAGCCGATCT 588
Qy      589 TCTAAATTTCAATCAAGTTTCACCTAGAGTTCTCAGGGAAGTTACCCCATCGCTCAATT 648
Db      589 TATAATCCCGCAAGTTTCAGCAAGTTTCTCTTGGAAAGTTATCTCTGCTCATATC 648
Qy      649 TCATACGAGTTAAACGTTCTAGGCTCTCGGGGACCGAGAGAATGGCTTCATCATGGAT 708
Db      649 ACATATGAGCTGATGTTCTTAGGATCCCGGGGACCAAGAAGATGCAATGCTTATGGAC 708
Qy      709 ACAATACCTATGAGCATCGTGGAGTCGCGAGGAGTAGTAGCTTCAACATCCATAAGCTCT 768
Db      709 ACAATACCTACAAGCACAATGAGGAGCTCAAGGAGTAGCTTCAGAACCATCAGAGTTTCCC 768
Qy      769 TTTTTCAGTCGGTTCATCACCAGTCTTTAGGTCCTCACTCAAAACCATTTGGCAGTAATAGT 828
Db      769 TTTACTCGGTACTCGGTCAACCTTTATCCAGTCTCAGTCAAAACCATTTACGCACTA----- 823
Qy      829 GCATCATGTAGCGACTCAGGCAACAACCTGGGAGATCCACATTTGGTGTGAGCAACAAA 888
Db      824 -----GCTCAAGCCACCTGAAAGAACACCATTAGTGTGAGCAACNAG 887
Qy      889 GCTCCACGGTGGATGAGCAGTTACGTTGCTGTGCTTAAATTTCCATGGTCCAGTCAACA 948
Db      868 ACACACGGTGGCAGGAGCAGTACGCTGCTGTGCTGTAATTTCCATGGCCGCTGTCA 927
Qy      949 GTGGCTTCGTTAAGAACTTTAGCTTGTGGAG-----TTAGTGAATGTGAAGCA 999
Db      928 GTAGCGTCAGTGAAGAACTTTACGCTCGTGGCAGCAGGAGCTAGCTGTGGCAGTGGCAG 987
Qy      1000 GGGCAGACATCTCAGA-----GGATCATCTCCAGTTTGGAAAGTTGGGAAG 1047
Db      988 GGAATGTACCGAGAGGCGAGCGAGCGGATTAATGTCAGTTTGGAAAGTTCGGGAAA 1047
Qy      1048 GACATGTTTACCATGGATTATGGATATCCGATTTCTGCGTTTCAAGCGTTTGTCTATCTGC 1107
Db      1048 GATATGTTTACGATGATATGATGATACCGATCTCAGCTTTCAGGCTTTTGGCAATTCG 1107
Qy      1108 CTGAGCAGTTTGAACACCAAGTTGCTGTGAA 1140
Db      1108 TTGAGCAGCTTTGAGACTAGAAATCGCTTGTGAA 1140

```

RESULT 8

BT013601 1781 bp mRNA linear PLN 11-MAY-2004
 LOCUS Lycopersicon esculentum clone 132366F, mRNA sequence.

DEFINITION

BT013601

ACCESSION

BT013601.1 GI:47105016

VERSION

FLI_CDNA.

SOURCE

Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 1781)

Kirkness, E.P., Wang, W. and Vazeille, A.

Submitted (11-MAY-2004) The Institute for Genomic Research, 9712

Medical Center Drive, Rockville, MD 20850, USA

Location/Qualifiers

1. 1781

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/db_xref="taxon:4081"

/clone="132366F"

/tissue_type="fruit"

/note="TZXB51"

FEATURES

source

ORIGIN

Query Match 30.8%; Score 350.6; DB 15; Length 1781;
 Best Local Similarity 59.2%; Pred. No. 6.1e-102;
 Matches 678; Conservative 0; Mismatches 449; Indels 18; Gaps 4;

```

Qy      11 GAAGTTTACTCCAGGAAATCGCGTCTAGGCCACACCGTGTAGTTTACGCCGCGCCCTCAA 70
Db      144 GCATTTCTCGGAAAGATTTGCTCCAGGTCAATAGGGTGGTTCAGGATTTACCGTGA 203
Qy      71 CCGCTAATAGTTTCAGACCCCTTTTCAGCTGTGCGAGTCCCGGAGGAGCTCTTAGAGAAA 130
Db      204 CTATCGATGCGTTGAGGCAAGTTGTTGGCCAACTGCGCCTGAGCTATTGAGAGATG 263
Qy      131 TCCGTATAGGGTTGAGACTGTTGACGGCGGCCGATTTGGCCGTTCGGCGGCAAAAGTGGTGG 190
Db      264 TCGTATAGGATTTGAGGAGTCGGA---CTCTGATTTGGCGTCCCAAGGAAAATTTGTTGG 320
Qy      191 CTTGTCCCGCGGTTTCTGCTAGCTGAGGATTTCTCAACAGGAGATTTGTAGCTGTTCTCTG 250
Db      321 CGTGCCTGGTGTTCGAGGATTTGAGGGAATCATGAAGAAGATTTGTTAAACACCTG 380
Qy      251 AATTCTCTCTAAATTTGACTTTCCCTATCTCCCTCAAGCAGTCTGTGTTCAAGAGATTTCTC 310
Db      381 AGTTAGCGGAAATTTGAGCTTCCCAATCTCTTTGAAGCAGCTGTGTTCAAGAGACAAAC 440
Qy      311 TAGTTCAATGCTTTATAAAACGTAATCGAAATACTCAATCGTATCATCTCTATCTCGGAT 370
Db      441 TTGTTCAATGTTTTTCATCAGAAAGAACCGCAGCTCACAGACATATCACCTTTTCTTAATT 500
Qy      371 TAACTACCTCTTTGACGATAACGGGAAGTTTCTCTGCTGCTTCTTAAGCTGAAGCGCG 430
Db      501 TAAATGAAGTTGTAAACGATGCGGTAAGTTCTCTTGTCTCGGAAGTGTAAAGGC 560
Qy      431 CAACTTGCATCTGATTAACATCATCTTTTGTGTTTCAGACGATATCTCAAAGAGAGCAACG 490
Db      561 CAACATACACAGATTAACATCATTTCTTTAAATCCGAAGTTGCTTCCAGSGTAGCAGCA 620
Qy      491 CGTATCTTGGGAGAAATGAGATCGAACTTCTCTTGGAAACAAATTCACGGTCTTTGATGTA 550
Db      621 CTTACATTTGGGAGCTTGAGATCCAACTTCTTGGGACCAAGTTCACTGTCTATGATGCAC 680
Qy      551 GTC---AGACCGGAGCAGCGAGATGACAGAGAGCGCTCTTCTAATTTTCATCA----- 601
Db      681 AACCATCTAATGACGAGCTTAAAGTCAGTAAATGTGCTCTACTAGTTTGTGGGATGA 740
Qy      602 ---AAGTTTCACTAGAGTTTCTCAGGGAAGTTTACCCCACTCGCTCACTTTCATACGAGT 658
Db      741 AACAGTCTCTCAAGAAATACCACTGGCAACTATCTGTAGTCTACATCTCGTATGAGT 800
Qy      659 TAAACGCTTTAGGCTCTCGGGGACCGAGAAAGATGCGTTGATCATGATGATGATCAATACCTA 718
Db      801 TGAATGCTCTGGGGGCTAGGGGGCCCGAGGAGATGCTGTGATATCATGAGCAATTCAG 860
Qy      719 TGAGCATCGTGGAGTCGCGAGGAGTAGTAGTTC---AACATCCATAAGCTCTTTTCCA 775
Db      861 CTTCTTTCTAATGAACTTGGGGGTTAGCCCTACGCAACAGAAATTTGTTCCAGTGAATG 920
Qy      776 GTCGGTCAATCACCAGCTTTTAGTGTCTCACTCAAAAACCATTTGCCAGTAAATAGTGCATCAT 835
Db      921 TAGATTTCTTCCCTCCCTCCATCTTTTAGATCAAAATCACTCGGTGGATTAATTTCT 980
Qy      836 GTAGCGACTCAGGCAACAACTGGGAGATCCACATTTGGTGTGAGCAACAAAGCTCCAC 895
Db      981 CATCCGGGCTTTTACCAGTCCAAAAGATGAAGTTCTAACTTTGAAGAAACAAAGCTCCTA 1040
Qy      896 GGTGGCATGAGCAGTTAGCTGTGCTGCTTAAATTTCCATGTTGAGTCAAGTCAAGTGGCTT 955
Db      1041 GGTGGCATGAAACAACTCCAAATGCTGTGCTTAACTTTCAATGCGCGGGTGAAGTGGCTT 1100
Qy      956 CGGTTAAGAACTTTAGCTTGTGGAGTGTAGTCACTGTGAAGCAGGCGACACATCTCAGA 1015
Db      1101 CTGTAAAAAATTTTACAGTAGTTGCTTCTCTTGGGATGGAGCTGGACAGAAATGAGA 1160

```

QY 1016 GGATCATACTCCAGTTTGGGAAGTTGGGAGGACATGTTTACCATGGATTATGATATC 1075
Db 1161 ATGTTATCTCCAAATTTGGGAAGTGGGAGAGANTTTGTTCCACTGGACTATCAGTATC 1220
QY 1076 CGATTTCTCGGTTTCAAGCGTTTGTCTATCTGCTGAGCAGATTTTGAACCAAGAAATTCCT 1135
Db 1221 CTATCTCGCATTTTCAGGCATTGCTATCTGCTCAGTAGCTTTGACACCAAAATCGCT 1280
QY 1136 GTGAA 1140
Db 1281 GCGAA 1285
RESULT 9
CQ804888
LOCUS CQ804888 1221 bp DNA linear PAT 10-MAY-2004
DEFINITION Sequence 1299 from Patent WO2004035798.
ACCESSION CQ804888
VERSION CQ804888.1 GI:47111078
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
Inze, D., de Veylder, L. and Vlieghe, K.
Identification of novel e2f target genes and use thereof
Patent: WO 2004035798-A 1299 29-APR-2004;
CropDesign N.V. (BE)
FEATURES
source 1..1221
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Query Match 30.1%; Score 342.6; DB 6; Length 1221;
Best Local Similarity 59.8%; Pred. No. 2.4e-99;
Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;
QY 32 GGTCTAGGGCCACACCGTGTAGTTTCACGCCGCCCTCAACCGCTAATAGTTCAGACCCCTT 91
Db 89 GATCCAGGCTCAACGTGTGTTTCAGATATCTTCTGTTCTGTGATGCTTTCAAGCAGA 148
QY 92 TCAGCTGTGCTGGAGCTCCCGAGGAGCTGCTTAGAGAAATCCCTGATTAGGGTTGAGACTG 151
Db 149 GCTGCTGGCTAGTATGCTCCGGAGCTCTCGAGAGATGTTCTTATGAGGATTGAGCAAT 208
QY 152 TTGACGGCGCGATGTCGCTCGCGCGAAACGTGGTGGCTTGTGCGCGGTTTGTGCTGA 211
Db 209 CCGAAGACAC- --TTGGCGCTCTAGGAAATAATGTTTCTTGGCTGTGTCTGCAGGA 265
QY 212 GCTGGAGGATCTCACCAAGGAGATGTAGCTCTCTGAATCTCTCTAAATTCACATT 271
Db 266 ACTGGCGAGAAATCGTCAAGAGATGCTCAGATCTCTGAGTCTCTAGCAAACTCACCT 325
QY 272 TCCTATCTCCCTCAAGCAGCTGTGCTCCAGAGATCTCTAGTTCATGCTTTATATAAAC 331
Db 326 TTCCTATCTCCCTCAACACCGCGGTCTCAGAGATCACTTGTTCATGCTATATTATGA 385
QY 332 GTAATCGAATATCAATGATATCTATCTATCTCGGATTA- --CTACCTCTTTGACGG 388
Db 386 GAAACCGCAGCAATCAAACTACTATCTATACCTCGGTTTAAACCAAGCAGCTTCAATG 445
QY 389 ATAAACGGAGTTTCTTCTGCTGCTTCTTAAGCTGAAGCGCAACTTGCACCTGATTACA 448
Db 446 ATGATGGAAGTTCTCTTCTTCTGCTGCAAGAGGTTTCGGAGGCCAACTTGCACCTACTA 505
QY 449 TCATCTCTTTGCGTTTCAGACGATATCTCAAGAGAGCAACGCGTATCTTTGGGAGATGA 508
Db 506 TCATCTCTTAACCTGGATGATGCTCTCTCGAGGAGCAATACCTATATCGGAAAGCTTA 565

QY 509 GATCGAATCTTCTTGGAAACAAATTTCAAGTCTTTTGATGGTAGTCAGACCGGAGCAG- -- 565
Db 566 GATCTAACTTTCTGGGACCAAGTTCACTGTCTATGACGCTCAGCCGAGAAATCTTGGA 625
QY 566 - -- 616
Db 626 CTCAGGTTTACCAAGAACCCGTTCAAGCAGACTTCTCAGTTTGAACAAGTGAGCCCGAGAA 685
QY 617 TTCCTCAGGAAGTTACCCCATCGCTCAATTTTCATACGAGTTAAACGTTCTTAGGCTCTC 676
Db 686 TTCCATCTGGCAACTATCTCTGTAGCACAATATCTATAGAGCTTAACGCTTTGGGTTCCA 745
QY 677 GGGGACCGAGGAAGTGGTTGTCATCGATACAAATACCTATGACATCGTGGAGTCGC 736
Db 746 GAGGACCGAGGAGGATGAGTGTGTCATGGATGCCATCCCTGCATCAGCTGTAGAACCTG 805
QY 737 GAGGAGTAGTAGTCTTCAACATCCATAAGCTCTTTTCCAGTC- -- -- -- -- -- -- 790
Db 806 GAGGAACAGCTCCAACCTCAGACGGAACTTGTCCATAGCAATCTTGATAGTTTCCCTCAT 865
QY 791 TCTTTAGTCTCAGTCAAAACCAATTTGGCAGGTAATAGTGCATCATGTAGGACTCAGGCA 850
Db 866 TCTCCTTCTTCAAGTTCGAAATCAATTCGTGCAGAGAGTCTCCCTTCTGCTCCATCATCTG 925
QY 851 ACAACCTGGAGATCCACCATTTGGTGTGAGCAACAAGCTCCAGCTCCGATGAGCAGT 910
Db 926 CTGCTCAGAGGAGGAGCTGTTGCTTCCGAAACCAAGGCCAGATGCGCAGAACAGC 985
QY 911 TACGTTGTGCTGTCTTAAATTTCCATGTGTCGAGTCACAGTGGCTTGGTTAAGAACTTTC 970
Db 986 TCCAGTGTGCTGTGCTCAACTTCAATGGGAGAGTCACAGTTGCTTCCGTCAAAAACCTTC 1045
QY 971 AGCTTGTGCGAGTGTAGTACTGTGAA- -- -- -- -- -- -- -- -- -- -- -- -- 1027
Db 1046 AGCTGTGTAGTCTCTCTCTGAGAAATGGACCTCGAGGACCTGAGCAGCAAGATGCTGAA 1105
QY 1028 AGTTTGGGAAGCTTGGGAGGACATGTTTACCATGATATGATATCGATTTCGCTGCGT 1087
Db 1106 AGTTTGGGAAGCTCGGAAAGATGTTTCAATGATATCAGTACCTTCTCTGCTTCTGCT 1165
QY 1088 TTCAAGCGTTTCTCTATCTGCTGAGCAGTTTGTGAAACCAAGAAATTCCTGTGAA 1140
Db 1166 TCCAGGCTTCACTTCCCTCAGCAGTTTCGACACCAAGATGATGCTGAA 1218
RESULT 10
AY045774
LOCUS Arabidopsis thaliana 1221 bp mRNA linear PLN 15-APR-2004
DEFINITION Arabidopsis thaliana tubby-like protein 3 mRNA, complete cds.
ACCESSION AY045774
VERSION AY045774.1 GI:27372513
KEYWORDS Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1221)
Lai, C.P., Lee, C.L., Chen, P.H., Wu, S.H., Yang, C.C. and Shaw, J.F.
Molecular Analyses of the Arabidopsis TUBBY-Like Protein Gene
Family
Plant Physiol. 134 (4), 1586-1597 (2004)
PUBMED 15064372
REFERENCE 2 (bases 1 to 1221)
AUTHORS Lai, C.P. and Shaw, J.F.
TITLE Cloning and characterization of cDNAs for tubby-like protein 3
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1221)
AUTHORS Lai, C.P. and Shaw, J.F.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Institute of Botany, Academia Sinica,
Nankang, Taipei, Taiwan 11529, Republic of China

QY	737	GAGGAGTAGTGGTTCACATCATTAAGCTCTTTTCCAGTC-----GGTCATCACCAG	790
DB	806	GAGGAACAGCTCCAACCTCAGACGGAACTTGTGCATAGCAATCTTGATAGTTTCCCTCAT	865
QY	791	TCTTTAGGCTCTCACTCAAAACCAATTCGCGCAGTAATAGTCATCATGTAGCGACTCAGGCA	850
DB	866	TCTCCTTCTTTTCAGGTGGAATCAATTCGTGCGAGAGAGTCTCCCTTCTGTCCATCATCTG	925
QY	851	ACAACTGGGAGATCCACCATTTGGTGTCTGAGCAACAAAGCTTCAACGGTGGCATGAGCAGT	910
DB	926	CTGCTCAGAAGGAAGGACTGCTTGTCTCGTGAATAACAAGCGCCAGATGGCAGCAACAGC	985
QY	911	TAGCTTCTCGTCTTAAATTTCCATGTGCTCAGTCAAGTGCCTTCGGTTTAAGAATTTC	970
DB	986	TCCAGTGTCTGGTGCCTCAACTTCAATGGGAGAGTCAAGTTGCTTCCGTCACAAAACCTTC	1045
QY	971	AGCTTGTGGCAGTTAGTGAAGTGTGAA---GCAGGGCAGACATCTGAGAGGATCATACTCC	1027
DB	1046	AGCTGTAGTGTCTCTGAGAAATGGACCTGCAGGACCTGAGCACGAAACGTTGATCTCC	1105
QY	1028	AGTTTGGGAAAGTTGGGAAGGACATGTTTACCATGATTATGATATCCGATTTCTGCGT	1087
DB	1106	AGTTTGGAAAAGTCGGAAAAGATGTGTTCACAAATGGAATTATCAGTACCCTATCTCTGCT	1165
QY	1088	TTCAAGCGTTTGTCTATCTGCTGAGCAGTTTGGAAACCAAGATTGCTCTGTGAA	1140
DB	1166	TCCAGGCGCTTTCACCATTTGCTTCAGCAGTTTCGACACCAAGATAGCATGTGAA	1218
RESULT 11			
LOCUS	AY096604	1252 bp	mRNA linear PLN 18-SEP-2002
DEFINITION	Arabidopsis thaliana putative tubby protein (At2g47900) mRNA, complete cds.		
ACCESSION	AY096604	1	GI:20465542
VERSION	AY096604.1		
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,		
	Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,		
AUTHORS	Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,		
	Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,		
	Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,		
	Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,		
	Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,		
	Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Arabidopsis Open Reading Frame (ORF) Clones		
	Unpublished		
JOURNAL	2 (bases 1 to 1252)		
	Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,		
REFERENCE	Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,		
	Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,		
AUTHORS	Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,		
	Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,		
	Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,		
	Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,		
	Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
	Direct Submission		
TITLE	Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan		
	Street, Albany, CA 94710, USA		
COMMENT	The RIKEN Genomic Sciences Center (GSC) members carried out the		
	collection and clustering of RAPL cDNAs (RAPL cDNA : RIKEN		
	Arabidopsis Full-length cDNA): Seki,M., Narusaka,M., Ishida,J.,		
	Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,		
	Hayashizaki,Y. and Shinozaki,K.		
	The Salk, Stanford, PGEC (SSP) Consortium members constructed and		
	sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada,K.,		

Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.	
Yamada, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinoraki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC) contributed equally to this work as PIs.	
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.	
FEATURES	Location/Qualifiers
source	1. .1252
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/chromosomes="2"
	/clone="U19776"
	/ecotype="Columbia"
	/note="This clone is in pUNI 51."
gene	1. .1252
CDS	/gene="At2g47900"
	1. .1221
	/gene="At2g47900"
	/codon_start=1
	/evidence=experimental
	/product="putative tubby protein"
	/protein_id="AA20254.1"
	/db_xref="GI:20465543"
	/translation="MSFKSLIQMRGELGISRKGFDPVFGVRSRSQRVVDTSVPVDAFKGSWASMPPELLDRLVLMRIEQLSDTPSRKNVSCAGVCRNWRREIVKIVRPELSKFTFTPLSLQPGRGLSVQYIMRNSQTYLYLGLNQAANDGKFLLAARFRPRTCTDYIISLNCDDVSRSNTYIIGKLSNFLTFTYDQNPNGQVTRTSRRLSKQVSPRIIPSGNPVAHISVELNLGSRGPRMOCVMDAIPASAVEPGGTAPTQT ELVSNLSDSPSPSPRSKISRAESLPSGSSAAOKEGLLVKLNKAPRHEOLOQWCLNFGNRTVASVKNFQLVAEPNGPAGPEHENVLQPGKVGKDVFTMDIYQYPISAQAF TICLSDFDTKIACE"
3' UTR	1222. .1252
	/gene="At2g47900"
ORIGIN	
Query Match	30.1%; Score 342.6; DB 15; Length 1252;
Best Local Similarity	59.8%; Pred. No. 2.4e-99;
Matches	677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;
QY	32 GGCTTAGGCCACCGCTGTAGTTTCACGCGCGCGCCTCAACCGCTAATAGTTTCAGACCCCTT 91
DB	89 GATCAGGCTCAACGCTGTGTTTCAGGATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 148
QY	92 TCAGCTGGTCGGAGCTCCCGAGAGAGCTGCTTAGAGAAATCCTGATTAGGTTGAGACTG 151
DB	149 GCTGCTGGGCTAGTATGCTCCGGAGCTCTCGAGATGTTCTTATGAGGATTTGAGCAAT 208
QY	152 TTGACGGCGCGATTGGCGCTCCGCGGGAACGCTGCTGCGCGGGTTGTCGTA 211
DB	209 CCGAAGACAC---TTGCGCGCTTAGGAAATAATGTTGTTCTTGGCTGGTGTCTGAGGA 265
QY	212 GCTGGAGGATTTCTCACCAGGAGATGTAGCTGTTCTGAATTTCTCTCTCTCTCTCTCTCT 271
DB	266 ACTGGGAGAAATCGTCAAGAGATGTCAGAGTTCTGAGCTTCTAGCAAACTCACT 325
QY	272 TCCTATCTCCCTCAAGCAGCTGGTCCAGAGATTTCTTAGTTCAATGCTTTATAAAC 331
DB	326 TTCTATCTCCCTCAAGACCGGGCTCTAGAGATCACTTGTCTAATGCTATATTATGA 385
QY	332 GTAATCGAATATCAATGCTATCACTCTATCTCGGATTA---CTACCTCTTTACCG 388
DB	386 GAAACCGCAGCAATCAAACTCTATCTATACCTCGGGTTAAACCAAGCAGCTTCAATG 445
QY	389 ATAAACGGAGTTTCTTCTGCTGCTCTAAGCTGAAGCGCGCAACTTGCACCTGATTACA 448

DB	446 ATGATGGAAGATTCTTCTTGCTGCAAGAGGTTTTCGGAGGCCAACTTTCGACTGACTACA 505
QY	449 TCATCTCTTTGGCTTCAGACGATATCTCAAAGAGAGCAACGCGTATCTCTTGGGAGATGA 508
DB	506 TCAUTCTCTTAACTGCGATGATGCTCTCGAGGAGCAATACCTATATTCGGAAGACTTA 565
QY	509 GATCGAACTTCTCTTGGAACAAAATTCACGGTCTTTTGATGGTGTAGTCAGACCGGAGCAG--- 565
DB	566 GATCTAACTTCTTGGGACCAAGTTCACTGTCTATGACGCTCAGCCGAGAACTCTGGAA 625
QY	566 -----CGAAGATGAGAGAGCCGCTCTTTCTAAATTTCTATCAAGTTTCACTTAGAG 616
DB	626 CTCAGGTTTACAGAACCCGTTCAAGCAGACTTCTCAGTTTGAACAAGAGTGAAGCCGAGAA 685
QY	617 TTCTCTCAGGAGAGTTTACCCCATCGCTCATTTTCATACGAGTTTAAACGCTTTAGGCTCTC 676
DB	686 TTCCATCTGGCAACTATCTCTGTAGCACAATCTCATATGAGCTTAAACGCTTTGGGTTC 745
QY	677 GGGGACCGAGAGAAATGCGTTGTCATCATGGATACAAATACCTATGAGCATCGTGGAGTCGC 736
DB	746 GAGGACCGGAGGATGAGTGTGTCATGGATGCCATCCCTGCATCAGCTGTAGAACCCTG 805
QY	737 GAGGAGTAGTAGCTTCAACATCCATTAAGCTCTTTTTCAGATC-----GGTCATCACAG 790
DB	806 GAGGAACAGCTCCAACTCAGACGGAACCTTGTCCATAGCAATCTTGATAGTTTCCCTCAT 865
QY	791 TCCTTTAGTCTCACTCAAAACCATTTGCGCAGTAATAGTGCATCATGTAGCAGCTCAGGCA 850
DB	866 TCTCTCTTTCAGGTCGAAATCAATTTGCTGAGAGAGTCTCCCTCTTGTGTCATCATCTG 925
QY	851 ACNACTCGGAGATCCACCATTTGGTGTGAGCAACAAAGCTCCACGGTGGCATGAGCACT 910
DB	926 CTGCTCAGAGAGAGAGCTGTTGCTTGAAACAAAGCGCCAGATGGCAGACAGC 985
QY	911 TACGTTGCTGGTCTTAAATTTCCAGTGTGAGTCAAGTGGCTTCGGTTTGAAGAACTTTC 970
DB	986 TCCAGTGTGCTGCTCAACTTCAATGGAGAGTCAAGTTGCTTCCGTCACAAAACCTTTC 1045
QY	971 AGCTTGTGCGAGTTAGTGAAGTGA---GCAGGCGAGACATCTGAGAGGATCATACTCC 1027
DB	1046 AGCTGTGTAGCTCTCTGAGAAATGGAACCTGCGAGGACCTGAGCAGCAAAACGTTGTTCTCC 1105
QY	1028 AGTTTGGGAAAGTTGGGAGGACATGTTTACCATGGATTATGGATATCCGATTTCTGCGT 1087
DB	1106 AGTTTGGGAAAGTCGGAAGAGATGTTTCAATGGATTATCAGTACCCTATCTGCTCT 1165
QY	1088 TTCAAGCGTTTGCTATCTGCTGAGAGGTTTGAACCAAGAAATTCCTGTA 1140
DB	1166 TCCAGGCTTCCACCATTTGCTTCCAGAGTTTCGACCAAGATAGCATGTGAA 1218
RESULT 12	
AY074273	
LOCUS	
DEFINITION	
Arabidopsis thaliana putative tubby protein (At2g47900) mRNA,	
complete cds.	
ACCESSION	
AY074273	
VERSION	
AY074273.1	
SOURCE	
FLI CDNA.	
ORGANISM	
Arabidopsis thaliana (thale cress)	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;	
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.	
1 (bases 1 to 1943)	
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,	
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,	
Carrinci, P., Chen, H., Cheuk, R., Hayaehizaki, Y., Ishida, J.,	
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,	
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,	
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinoraki, K.,	
Davis, R.W., Ecker, J.R. and Theologis, A.	
Arabidopsis Full Length cDNA Clones	
TITLE	

JOURNAL REFERENCE AUTHORS	Unpublished 2 (bases 1 to 1943) Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, J.C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. Direct Submission Submitted (11-JAN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.	Db
TITLE JOURNAL COMMENT	The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yanada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A. Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC) contributed equally to this work as PIs. Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank. Location/Qualifiers 1. .1943 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="2" /clone="RAFL09-89-B17 (R19776)" /ecotype="Columbia" /note="This clone is in a modified pBluescript vector (pLC-1) as a BamHI/XhoI insert." 1. .1943 /gene="At2g47900" 1. .454 /gene="At2g47900" 455. .1675 /gene="At2g47900" /codon_start=1 /evidence=experimental /product="putative tubby protein" /protein_id="AAL66970.1" /db_xref="GI:18377640" /translation="MSFKSLIQDMRGELGISRKGFDFVFGVGRSRQRVQDTSPVPDAFKQSCAMPELLRDVLMRIESEDTPSRKXNVSCAGVCRNWRRETVKIVRYPELSKLTFTSLKPGPRGSLVQCYIMRNSQTYLYILGNQASNDKGFLLAAKRFRRPTCTDYIISLNCDDVSRSGTYIGKLSNFGTKFTVYDAQPTNPQVTRTRSSRLSLKQVSPRIPSGNPVAHISVELNVLGSRPRMQCVMDALPASAVPGTAPTQTELHVNLDSPFSPFRSKSIRAESLPSGSSNAQKEGLLVLNKAPRWHEQLQWCLNFGRTVAVSKNFQVLAAPENGPAGPEHNVILFGKGVDFNDYQYPSAFQAF TICLSFDTKIACE" 1676. .1943 /gene="At2g47900"	Qy
3' UTR		Qy
ORIGIN	Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;	Qy
5' UTR		Qy
CDS		Qy
gene		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best Local Similarity 59.8%; Pred. No. 2.5e-99; Matches 677; Conservative 0; Mismatches 429; Indels 27; Gaps 5;		Qy
5' UTR		Qy
CDS		Qy
3' UTR		Qy
ORIGIN		Qy
Query Match 30.1%; Score 342.6; DB 15; Length 1943; Best		

RESULT 13

AK071159 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023078015, full insert sequence.
 AK071159 ACCESSION
 VERSION AK071159.1 GI:32981182
 FLY CDNA; CAP trapper.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurokawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
 Japoneica rice
 Collection, mapping, and annotation of over 28,000 cDNA clones from Japonica rice
 Science 301 (5631), 376-379 (2003)

TITLE

Japoneica rice
 Science 301 (5631), 376-379 (2003)

JOURNAL

PUBMED 12869764

REFERENCE

AUTHORS 2 (bases 1 to 1687)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Maeda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
 Direct Submission

TITLE

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp)
 Tel: 81-29-838-7007, Fax: 81-29-838-7007

JOURNAL

COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Maeda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
 and Genome Science Laboratory in Riken Genomic Sciences Center
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
 Location/Qualifiers
 1..1687
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J023078015"

ORIGIN

Query Match 29.5%; Score 336.2; DB 15; Length 1687;
 Best Local Similarity 59.5%; Pred. No. 3e-97;
 Matches 632; Conservative 0; Mismatches 413; Indels 18; Gaps 3;
 QY 93 CAGCTGTCGGAGCTCCCGAGGAGCTGTTAGAGAAATCTGTAGGTTGAGACTGT 152
 DB 350 CTGCTGCGAGCTGCGCGCGAGCTCTCTCGGGAGGTGCTGGCGAGGTGGAGGATC 409
 QY 153 TGACCGCGCGGATTGGCGGTGCGCGGAACCTGCTGCGCGGCTTGGCGGCTTGTCTAG 212
 DB 410 GGAGGGGTG---GGCGCGCGCGCGGACGCTGGTGGCGCGCGCGGCTGCGCGGAG 466
 QY 213 CTGGAGGATTTCTCACCAAGGAGATTGTAGCTGTTCTGTAATTTCTCTCTAAATTTGACTTT 272
 DB 467 CTGGAGGGGATCTGTCAGGAGATCTGCGCACCCCGGAGGCGTGGGAACTCACCTT 526
 QY 273 CCTATTCCTCCCAAGCAGTCTGGTCCAAAGAGATCTCTAGTTCATGCTTTATAAAACG 332
 DB 527 CCCAATCTCTCTCAAGCAGCTGGCGCGAGGATGCTCTCTATGAAATGTTTCATTGTAAG 586
 QY 333 TAAATCGAATATCTCAATCGTATCATCTCTATCTCGGATTAACCTCTCTTTCGCGGATAA 392
 DB 587 GAACCGGACCATCTAGACATATTTACTCTGACATTTGGACTGATGATGATGATGATGATGA 646
 QY 393 CGGGAGGTTTCTCTTCTGCTGCTTCTAAAGCTGAAGCGCGCAACTTGCATGATTACATCAT 452
 DB 647 TGGGAAGTTCTCTACTTCTGTCACGCAATATGTCGTAGGACCACTGACCTGAGTACCTGAT 706
 QY 453 CTCTTTTGGCTTCAGACGATATCTCAAGAGAGAACCGGTATCTTGGGAGGATGATGATC 512
 DB 707 TTTCGCTTGATGAAACGATATATCAAGAGGACTGATAGTTATGTTGGCAAACTAAGATC 766
 QY 513 GAATCTCTTGGCAAAATTTCAAGGTTCCGGTCTTTTCATGTTGATGTCAGACCGGAGCAG- 565
 DB 767 TAACTCTTTGGGAACCAATTTACTATCTATGATGCCCATCCACCTTATGCTGGAGATGT 826
 QY 566 -----CGAAGATGCGAAGAGCGGCTCTTCTTAATTTTCATCAAGTTTTCACCTAGAGTTCC 620
 DB 827 GATTTCAAGGGTCTAGCTGCGAGGTGATGTTGGTTCAAAACCATTTATCCCCAGGATACC 886
 QY 621 TCAGGGAAGTTACCCCATCGCTCAATTTTCATACAGGATTAACGCTTTTAGGCTCTCGGGG 680
 DB 887 TGTCTGGGAATTTTCCAGTCTCATATATTTTCATATGAATTTGAACGTTTTGGGTTCTTAGGG 946

QY 681 ACCGAGAAGATCGTTCATCATGATACATACCTATGAGCATCGTGGAGTCGGGAGG 740
 Db 947 CCGGAGAAGATCGTTCATCATGATACATACCTATGAGCATCGTGGAGTCGGGAGG 1006
 QY 741 AGTAGTAGCTTCAACATCCATAGCTCTTTTCCAGTCGGTCTATCACCAGTCTTTAGGTC 800
 Db 1007 AACAGCTCCGACGCGAGACTGAAATTTCTTTGAGTTACCATGAATCTTTCACATCAATTC 1066
 QY 801 TCACTCAAAACATTCGGCAGTAAATAGTGCATCATGATAGGATCAGGACAACTCGG 860
 Db 1067 CTTTTTCAAAATCAAAATCAGTTTCGAGCGAATTAATTCATGCTTCTTACTCACTCAAAA 1126
 QY 861 AGATCACCATTGGTCTCGACCAACAAAGCTCCACGCTGSCATGACGAGTTAGCTTGCTG 920
 Db 1127 TGGAAACAACATGGTGTGGAAGAACAAAGTCCCGGAGGTGGCAGCAATCTGCGAGTGCTG 1186
 QY 921 GTGCTTAAATTTCCATGTCGAGTCACAGTGGCTTCGGTTAAAGAACTTTTCAGCTTCGTGC 980
 Db 1187 GTGCTTAAATTTCCATGTCGAGTCACAGTGGCTTCGGTTAAAGAACTTTTCAGCTTCGTGC 1246
 QY 981 AGTTAGTACTGTGA---AGCAGGGCAGACATCTGAGAGATCATACTCAGTTTGGGAA 1037
 Db 1247 TTCAGATCAGAGTAACCCCAACCAACAGGAAACAGATGATGATCTCCAGTTTGGTAA 1306
 QY 1038 AGTTGGAGGAGCATGTTTACCATGATATGATATCCGATTCGCTTTCAAGGTT 1097
 Db 1307 AGTTGGAGGAGCATGTTTACCATGATATGATATCCGATTCGCTTTCAAGGTT 1366
 QY 1098 TGTATCTGCTGAGCAGTTTGAACACCAAGAAATGCTGTGAA 1140
 Db 1367 TGGATATGTCGAGCAGCTTTGACACCAAAATAGGCTGGAA 1409

RESULT 14

AK060587
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:001-024-B07, full insert sequence.

ACCESSION AK060587

VERSION AK060587.1 GI:32970605

KEYWORDS FLI CDNA; oligo-capping.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Eurya sativa (japonica cultivar-group)

Biological processes: Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otono, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurotsaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Otsu, M., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
 Science 301 (5631), 376-379 (2003)
 12869764

2 (bases 1 to 1498)

ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., CARNINCI, P., DOI, K., FUJIMURA, T., FUKUDA, S., HANAGAKI, T., HARA, A., HASHIZUME, W., HAYASHIDA, K., HAYASHIZAKI, Y., HAYATSU, N., HIRAMOTO, K., HIRAKAWA, T., HORI, F., HOTTA, I., IIDA, J., IIDA, Y., IKEDA, R., IMAMURA, K.,

TITLE
JOURNAL

Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : http://cdna01.dna.affrc.go.jp/cDNA/

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, I., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers

1..1948
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="001-024-B07"

ORIGIN

Query Match 29.5%; Score 336.2; DB 15; Length 1948;
 Best Local Similarity 59.5%; Pred. No. 3e-97;
 Matches 632; Conservative 0; Mismatches 413; Indels 18; Gaps 3;
 QY 93 CAGCTGCTCGGAGCTCCCGAGGAGCTGTTAGAGAAATCTGATAGGCTTGGAGACTGT 152
 Db 405 CTGCTGGACGCGAGCTCCCGCGAGCTCTCGGGAGGTGCTGGCGGGTGGAGGATC 464
 QY 153 TGACGCGCGCGATTTGGCCCGTCCGCGGCGAAACGTTGGTGGCTTGTCCCGCGCTTTGT 212
 Db 465 GGAGGGGTG---GTGCGCGCGCGCGCGACGTTGGTGGCTGGCGCGCGCTCTGCCGAG 521
 QY 213 CTGGAGGATTTCTCACCAAGAGATTTGATGTTCTTCTGAATTTCTCTTAATTTGACTTT 272

Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, P.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. .2102
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033091L01"

FEATURES

source

23.1%; Score 262.8; DB 15; Length 2102;
Best Local Similarity 58.3%; Pred. No. 2.5e-73;
Matches 506; Conservative 0; Mismatches 347; Indels 15; Gaps 2;

ORIGIN

Query Match 23.1%; Score 262.8; DB 15; Length 2102;
Best Local Similarity 58.3%; Pred. No. 2.5e-73;
Matches 506; Conservative 0; Mismatches 347; Indels 15; Gaps 2;
QY 288 GCAGTCTGCTCAAGAGATTCTCTAGTTCATAGCTTTTATAAAGCGTAATCGAAATACTCA 347
DB 24 GCAGCTGCCCCAAGAGACGGCACTCTTAATGTTTCATAGAGAAACGGGACTACTCA 83
QY 348 ATCGTATCATCTCTATCTCGAATTAACTACCTCTTTGACGGATAACGGGAAGTTCTTCT 407
DB 84 GACATATTATCTGTATATTGGGCTCACAGAAGCACCTGCGGATGATGGGAAGTTCTTACT 143
QY 408 TGCTGCTTCTAGCTGAAGCGGCACTTGGCACTGATTACATCATCTCTTTGGTTTCA 467
DB 144 TCGTGCAACCAATGACGAAGCCCATGTCACAGACTACCTGATTTCCTTGACATGAG 203
QY 468 CGATATCTCAAAGAGAACGCGGTATCTTGGGAGAAATGAGATCGAACTTCTCTGGAAC 527
DB 204 TGTATGTCAAGGGGAGGACACCTATATTGGCAAGCTAAGGTCAAACTTTCTCGAAC 263
QY 528 AAAATTCACGGTCTTTGATG-----GTAGTCAGACCGGAGCAGCGAAGATGCA 575
DB 264 AAAGTTTACGGTCTATGATGCCATCCACATATGATGAGCTGTTGTCTCAAAAAGTCG 323
QY 576 GAAGAGCCGCTCTTCTAATTTTCATCAAGTTTCACCTAGAGTTCTCTCAGGGAAGTTACCC 635
DB 324 GTCTGCACGTGTGTTGGTCTGAACCAAGTCTCCCGAGAGTTCTTCCCGGGAATTTACCC 383
QY 636 CATCGCTCACATTTTCATACGAGTTAAACGTTCTTAGGCTCTCGGGGACCGAGAAGATGCG 695
DB 384 AGTTTCACATATTTCTACGAATTTGAATGTTTGGGTGCAAGAGTCCCAAGAGGATGAA 443
QY 696 TTGCATCATGGATACAAATACCTATAGCATCTGTTGGAGTCGCGAGGAGTAGTAGCTTCAAC 755
DB 444 CTGCATTATGGATTCCATCCCCACATCAGCTGTTCAAGAGGGTGGGAAAGCTCCGACACA 503
QY 756 ATCCATAAGCTCTTTTCCAGTCGGTTCATCAACAGTCTTTAGGTCTCATCTCAAAACCAT 815
DB 504 GACTGAATTTCTCTCAGTGGGCTCGACTCTTTCCCATCAATTTCAATTTCTCAGATCAAA 563
QY 816 GCGCAGTAAATAGTCATCATGTAGCGACTCAGGCAACACCTGGGAGATCCACCATTTGCT 875
DB 564 ATCAGCTCGAATAGATAGTGCAACTTTCAAAATATCCACTCAGAAGGAAGAAAGCTGGT 623
QY 876 GCTGAGCAAAAGCTCCACGGTGGCATGAGCAGATTACGTTGCTGTGCTTAAATTTCCA 935
DB 624 TCTGAAGATAAATCTCTAGTGGGATGAGCACTGCAATGTTGGTGGCTCAACTTCCG 683
QY 936 TGGTCGAGTCAAGTGGCTTCCGTTAAGAACTTTTCAGCTTTGGTGGCAGTTAGTGAAGTGG- 994
DB 684 TGGACGGGTTACTGTTCCTCAGTGAAGAACTTTTCAGCTGGTGGCTCTGATGAAATGG 743
QY 995 --AAGCAGGGGACACATCTGAGAGGATCATCTCCAGTTTGGGAAGTTGGGAAGGACAT 1052
DB 744 GCCAACTAACCGAGGACGACGAAGGTGATTTCCAGTTTGGGAAGATTTGGGAAGGACTT 803

QY 1053 GTTTACCATGGATTATGGATATCCGATTTCGGGTTTCAAGCGTTTGCTATCTGCTGAG 1112
DB 804 GTTCCACATGGACTACCGTTATCCGATATCAGCATTTCAATCATTTGCAATCTGTCTGAG 863
QY 1113 CAGTTTTGAAACACAGAAATTCGCTGTGAA 1140
DB 864 CAGTTTTGATACCAAAATAGCCTGTGAA 891

Search completed: December 26, 2005, 05:22:21
Job time : 6060 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 23:20:52 ; Search time 13 Seconds
(without alignments)
208.498 Million cell updates/sec

Title: US-10-763-042-9

Perfect score: 1962

Sequence: 1 WTRSLQEMRSPRHVHA.....SAFQAFICLSFFETRIACE 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 54001 seqs, 7132810 residues

Total number of hits satisfying chosen parameters: 54001

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pap:*
- 2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pap:*
- 3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pap:*
- 4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pap:*
- 5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pap:*
- 6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pap:*
- 7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pap:*
- 8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.5	4.3	865	US-11-109-156-4	Sequence 4, Appli
2	82.5	4.2	761	US-10-485-517-252	Sequence 252, App
3	82	4.2	651	US-11-172-145-10	Sequence 10, Appl
4	80.5	4.1	581	US-11-045-802-30	Sequence 30, Appl
5	80.5	4.1	617	US-11-172-145-8	Sequence 8, Appli
6	80	4.1	483	US-11-137-465-40	Sequence 40, Appl
7	80	4.1	605	US-11-137-465-41	Sequence 41, Appl
8	80	4.1	1075	US-11-174-150-34	Sequence 34, Appl
9	80	4.1	1114	US-11-174-150-35	Sequence 35, Appl
10	79.5	4.1	615	US-11-172-145-6	Sequence 6, Appli
11	79.5	4.1	750	US-11-070-627-2	Sequence 2, Appli
12	79	4.0	584	US-11-045-802-31	Sequence 31, Appl
13	79	4.0	841	US-10-770-726-88	Sequence 88, Appl
14	79	4.0	1897	US-10-821-234-1635	Sequence 1635, Ap
15	79	4.0	1907	US-11-000-463-250	Sequence 250, App
16	77	3.9	521	US-11-091-914-2	Sequence 2, Appli
17	76	3.9	437	US-10-632-150-54	Sequence 54, Appl
18	76	3.9	437	US-11-073-457-54	Sequence 54, Appl
19	76	3.9	437	US-11-073-460-54	Sequence 54, Appl
20	75	3.8	244	US-10-510-386-110	Sequence 110, App
21	75	3.8	826	US-10-821-234-1048	Sequence 1048, Ap
22	74	3.8	397	US-10-467-657-3814	Sequence 3814, Ap
23	73.5	3.7	708	US-11-174-150-25	Sequence 25, Appl
24	73.5	3.7	736	US-11-174-150-26	Sequence 26, Appl
25	73	3.7	415	US-11-467-657-7774	Sequence 7774, Ap

26	72.5	3.7	610	7	US-11-194-246-312	Sequence 312, App
27	72.5	3.7	736	7	US-11-078-189-9	Sequence 9, Appli
28	72.5	3.7	750	7	US-11-070-627-3	Sequence 3, Appli
29	72.5	3.7	1142	7	US-11-044-051-73	Sequence 73, Appl
30	72	3.7	765	6	US-10-821-234-1164	Sequence 1164, Ap
31	72	3.7	934	7	US-11-108-864-19	Sequence 19, Appl
32	72	3.7	945	6	US-10-131-826A-146	Sequence 146, App
33	71.5	3.6	216	7	US-11-055-822-58	Sequence 58, Appl
34	71	3.6	786	7	US-11-070-627-9	Sequence 9, Appli
35	71	3.6	805	6	US-10-467-657-8246	Sequence 8246, Ap
36	70.5	3.6	241	7	US-11-059-275-2	Sequence 2, Appli
37	70.5	3.6	798	7	US-11-107-028-2	Sequence 2, Appli
38	70.5	3.6	881	6	US-10-623-155-430	Sequence 430, App
39	70.5	3.6	920	6	US-10-623-155-357	Sequence 357, App
40	70.5	3.6	943	6	US-10-623-155-161	Sequence 161, App
41	70.5	3.6	1176	6	US-10-821-234-897	Sequence 897, App
42	70	3.6	428	6	US-10-632-150-6	Sequence 6, Appli
43	70	3.6	428	7	US-11-073-457-6	Sequence 6, Appli
44	70	3.6	428	7	US-11-073-460-6	Sequence 6, Appli
45	69.5	3.5	176	6	US-10-467-657-174	Sequence 174, App

ALIGNMENTS

RESULT 1

US-11-109-156-4
; Sequence 4, Application US/11109156
; Publication NO. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; TITLE OF INVENTION: PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-4


```
; Sequence 30, Application US/11045802
; Publication No. US20050257289A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Heientjaris, Tim
; APPLICANT: Lowe, Keith
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell
; APPLICANT: Zheng, Peizhong
; TITLE OF INVENTION: AP2 Domain Transcription Factor ODP2 (Ovule Development Protein 2)
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: 035718/286074
; CURRENT APPLICATION NUMBER: US/11/045,802
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 60/541,122
; PRIOR FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-045-802-30

Query Match      4.1%; Score 80.5; DB 7; Length 581;
Best Local Similarity 19.2%; Pred. No. 2.8;
Matches 61; Conservative 40; Mismatches 87; Indels 129; Gaps 14;

QY 50 TVDGDMPRRNNVACAGVCRSWRLTKRIIVAVPEFSSKLTTPISLKQSPRDSLQV--- 106
DB 171 TSDSNYNNDVVQ-----EKTIVDVETPKT-----IESFGRTSIYRGVT 215

QY 107 -----C-----FIKRNRTQSYHL-----YLGTLTSLTDN--- 131
DB 216 RHRWTGRYEAHLWDNSCKREGQTRKGGYDKEEKAAAYDLAALKYWGTTT--TTNFP 273

QY 132 -KFLLAASKLRATCTDYIISLR--SDDISKESNAY-----LGRMSN- 172
DB 274 LSEYEKEVEEMKHMTRQEVASURKSGFSRGASIRYGVTRHHQGRWOAIRGVAGNK 333

QY 173 --FLGTFTVFDGSGTGAARKQKRSNFIKVSPPVQGSYPIAHISYELNVLGSRGPRR 230
DB 334 DLYLGT-----FTQBERAE-----AYDIAAKF----- 357

QY 231 MRCIMDTPMSIVSRGVASTSISFSRSRSPVFRSHKPLRSNASCSDSGNNIGDPP 290
DB 358 -----RGLSAVTFDMNRYNVAILESPLIGSSAKRLKDVNNPV--PA 400

QY 291 LVLSNKAPRWHQRLQW 307
DB 401 MMISNNVSSANNVSGW 417

RESULT 5
US-11-172-145-8
; Sequence 8, Application US/11172145
; Publication No. US20050272696A1
; GENERAL INFORMATION:
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: METHODS OF SELECTIVELY TREATING DISEASES WITH SPECIFIC
; TITLE OF INVENTION: GLYCOSAMINOGLYCAN POLYMERS
; FILE REFERENCE: 3554.104
; CURRENT APPLICATION NUMBER: US/11/172,145
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/584,442
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 10/642,248
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: 60/404,356
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/479,432
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: 60/491,362
```

```
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 10/195,908
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/107,929
; PRIOR FILING DATE: 1998-11-11
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-11-172-145-8

Query Match      4.1%; Score 80.5; DB 7; Length 617;
Best Local Similarity 21.4%; Pred. No. 3;
Matches 55; Conservative 30; Mismatches 75; Indels 97; Gaps 14;

QY 7 LOEMSRPRHVVHAAASTANSSDPFSSWSELPEELLREILIRVETVDGGDWPGR-RNVVAC 65
DB 330 LKEIFSFP--IHDALPTISKEMSKLSNPKIP-----VYINICSI-----PSRIKQLQYT 376

QY 66 AGV-----CRSMRLTKRIIVAVPEFSSKLTTPISLKQSPRDSLQVCFIKRNRTQSYHLY 121
DB 377 IGVLNQCDHPHYLDGYPEVDFIKKL-----GNKATVINC-----QNKKN----- 417

QY 122 LGLTSLTDNGKFLAASKLR-----ATC-----TDYL-----ISL 153
DB 418 ----ESIRNGKFILLEKLIKENDGYITCDDDIRYPADYINTWIKKINKYNDKAAIGL 473

QY 154 RSDDISKRSNAYLGRMR-----SNEFLT-----KFTVFDGSQTGA 188
DB 474 HGVIFPSRVNKFSSDRIVYNFQKPLENDTAVNLTGTVAFRVSIFNKFSLSDFEHPGM 533

QY 189 AKMQKS---RSSNFIKV 202
DB 534 VDIYFSILCKKNILQV 550

RESULT 6
US-11-137-465-40
; Sequence 40, Application US/11137465
; Publication No. US2005025558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, P.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
```



```
QY 242 -----IVESRGVASTSISSFSRSPVFRSHSKPLR-----SNSASC 279
Db 988 LLVTDPTFCRMGSGPEILTLTVASAHNASVKCDMRSQKRRRPPDSCHAFHPEENAQC 1047
QY 280 SDSGNNLGDPLVL 293
Db 1048 GGASDTSASPLLL 1061

RESULT 9
US-11-174-150-35
; Sequence 35, Application US/11174150
; Publication No. US20050260714A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, Zhaoying
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50022
; CURRENT APPLICATION NUMBER: US/11/174,150
; CURRENT FILING DATE: 2005-07-01
; PRIOR APPLICATION NUMBER: US/10/257,174
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US01/11797
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/196,603
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/199,417
; PRIOR FILING DATE: 2000-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 1114
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-150-35

Query Match 4.1%; Score 80; DB 7; Length 1114;
Best Local Similarity 19.7%; Pred. No. 7.6;
Matches 62; Conservative 43; Mismatches 141; Indels 68; Gaps 11;

QY 33 WSELPELRLIRIVETVGGWGRNNVACAGVCRWRILTKIVAVPEFSSKLTTP 92
Db 802 WAEGPESAGEPMVVTASTAVAVTVDRKTAIAAAAGVQMKLEFLQRFMAATROCSTVDGP 861
QY 93 ISLKQSGPRDSLVOCFI-----KRNNTQSYHLVIG-----LTTSLTDNGKFLLA 137
Db 862 CT---QSCSDSDLDCTVIDNNGFILSKRSRETGR---FLGVDGAVLTQLLSMGVYF--- 912
QY 138 ASKLRKATCTDYIISLR-----SDDISKRSNAVLGRMRSNFLTGTFTVFDGSGTGA 189
Db 913 ----SQTWYDYQAMKPSHHHSAQPLVPSIFATRTATRLQLLVFLLENSWGWSW 968
QY 190 KMQKSSSNFIKVSRRVP-QGSYPIAHISY-----ELNVLGSRGPRRMCIMDTIPMS-- 241
Db 969 YDRGAZAHKHKQDPLQPCDTEYFV--FVYQPAIREANGIVECGPCQKVFVQQIIPNSNL 1026
QY 242 -----IVESRGVASTSISSFSRSPVFRSHSKPLR-----SNSASC 279
Db 1027 LLVTDPTFCRMGSGPEILTLTVASAHNASVKCDMRSQKRRRPPDSCHAFHPEENAQC 1086
QY 280 SDSGNNLGDPLVL 293
Db 1087 GGASDTSASPLLL 1100

RESULT 10
US-11-172-145-6
; Sequence 6, Application US/11172145
; Publication No. US20050272696A1
; GENERAL INFORMATION:
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: METHODS OF SELECTIVELY TREATING DISEASES WITH SPECIFIC
; FILE REFERENCE: 3554.104
; CURRENT APPLICATION NUMBER: US/11/172,145
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/584,442
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 10/642,248
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: 60/404,356
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/479,432
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 60/491,362
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 10/195,908
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/107,929
; PRIOR FILING DATE: 1998-11-11
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 615
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-11-172-145-6

Query Match 4.1%; Score 79.5; DB 7; Length 615;
Best Local Similarity 21.0%; Pred. No. 3.8;
Matches 54; Conservative 30; Mismatches 76; Indels 97; Gaps 13;

QY 7 LQEMSRPHRVVHAASSTNSDPFSSWSELPELRLIRIVETVGGWMPSR-RNVVAC 65
Db 328 LKEIFSPR--IHDALPISKEMSKLSNPKIP-----VVINICSI-----PSRIKQLQYT 374
QY 66 AGV-----CRSWILLTKEIVAVPEFSSKLTFTPISLKQSGPRDSLVOCFIKRNNTQSHLY 121
Db 375 IGVLRNQCDHFHYLDGVEYDFIKCL-----GNKATVINC-----QNK----- 415
QY 122 LGITTSLTONGKFLAASKLR-----ATCTDYI-----ISL 153
Db 416 ----ESIRONGKFLILEKLIKENKQGYITCDDDIRYPADYNTMTIKKYNKAAIGL 471
QY 154 RSDDISKRSNAYLGRMR-----SNFLGT-----KFTVFDGSGTGA 188
Db 472 HGVIPEFRVNYKFSSDRIVVNFQKPLENDTAVNLTGTVAPRVSIFNKFSLDSDFEHPGM 531
QY 189 AKMQKS-----RSSNFIKV 202
Db 532 VDIYFSILCKGNILQV 548

RESULT 11
US-11-070-627-2
; Sequence 2, Application US/11070627
; Publication No. US20050271625A1
; GENERAL INFORMATION:
; APPLICANT: Nash, Kevin R.
; APPLICANT: Burger, Corinna
; TITLE OF INVENTION: RAAV-NEPRILYSIN COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 36689.8
; CURRENT APPLICATION NUMBER: US/11/070,627
; CURRENT FILING DATE: 2005-03-02
; PRIOR APPLICATION NUMBER: 60/549,399
; PRIOR FILING DATE: 2004-03-02
```

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-11-070-627-2

Query Match 4.1%; Score 79.5; DB 7; Length 750;
Best Local Similarity 25.0%; Pred. No. 4.9; 28; Indels 35; Gaps 4;

Matches 26; Conservative 15; Mismatches 15; Indels 35; Gaps 4;

QY 30 PFSWSLPEELREILIRVETVDGDPGRNNVACAGVCRSWRLITKEIVAPFSSKL 89

Db 320 PFSWSNFTWEIMSTVNINQNEE-----EYVVAPEYLTKL 355

QY 90 TFPISLKQSGPRDSLVOCFIKRNRNTQSYHLVGLTTSITDNGK 133

Db 356 K-PILTKYS-PRDL-----QNLMSWRFIMDLVSSLSRNYK 388

RESULT 12

US-11-045-802-31

; Sequence 31, Application US/11045802

; Publication No. US20050257289A1

; GENERAL INFORMATION:

; APPLICANT: Gordon-Kamm, William

; APPLICANT: Heientjaris, Tim

; APPLICANT: Lowe, Keith

; APPLICANT: Shen, Bo

; APPLICANT: Tarczynski, Mitchell

; APPLICANT: Zheng, Peizhong

; TITLE OF INVENTION: AP2 domain Transcription Factor ODP2 (Ovule Development Protein 2

; FILE OF INVENTION: and Methods of Use

; FILE REFERENCE: 035718/286074

; CURRENT APPLICATION NUMBER: US/11/045,802

; CURRENT FILING DATE: 2005-01-28

; PRIOR APPLICATION NUMBER: 60/541,122

; PRIOR FILING DATE: 2004-02-02

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 31

; LENGTH: 584

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-11-045-802-31

Query Match 4.0%; Score 79; DB 7; Length 584;

Best Local Similarity 19.4%; Pred. No. 3.9;

Matches 62; Conservative 39; Mismatches 88; Indels 130; Gaps 13;

QY 50 TVDGDWPSRRNVACAGVCRSWRLITKEIVAPFSSKLTFPISLKQSGPRDSLVOCFI 109

Db 171 TSDSNYNNDVVQ-----EKTIVDVETTPKKT-----IESFGQRTSIYR-GV 214

QY 110 KRNRTQSY--HL-----YLG-----LTTSL 128

Db 215 TRHRWTGRYEAHLWDNSCKREGQTRGRQVYLGVDYKKEEKAARAYDLAALKYWGPTTTN 274

QY 129 TONGKFLAASKLKRACTCTDYIISLR--SDDISKRSNAY-----LGRMRS 171

Db 275 FPLSEYEKEVEEMKHTQRYVASLRKSSGFRGASIIYRGVTRHHGHGWQARIGRVAG 334

QY 172 N----FLGTFTVFDGQTGAANKQKRSNFIKVSFVPGQSGYPIAHISYELNVLSRGP 228

Db 335 NKDLYLGT-----FGTQEEAAE-----AYDIAAIFK----- 360

QY 229 RMRRCIMDTIPMSIVSRGVVASTSISFSSRSPPVFRSHSKPLRSNSASCSDSGNNLGD 288

Db 361 -----RGLSAVTNFDNMRYNKAILESPLSPIGSAKRLKDVNNPV-- 401

QY 289 PPLVLSNKAPRWHEQLRCW 307

Db 402 PAMMISNNVSESANNVSCW 420

RESULT 13

US-10-770-726-88

; Sequence 88, Application US/10770726

; Publication No. US20050266409A1

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Brown, Eugene

; APPLICANT: Liu, Wei

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING

; FILE OF INVENTION: CANCERS

; FILE REFERENCE: AM101079 (031896-010000)

; CURRENT APPLICATION NUMBER: US/10/770,726

; CURRENT FILING DATE: 2004-02-04

; NUMBER OF SEQ ID NOS: 48640

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 88

; LENGTH: 841

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-770-726-88

Query Match 4.0%; Score 79; DB 6; Length 841;

Best Local Similarity 19.6%; Pred. No. 6.5;

Matches 90; Conservative 62; Mismatches 175; Indels 132; Gaps 21;

QY 4 RSLQEMRSR---PHRVVHAAASTANSDDPFSWSLPEELL--REILIRVETVDG--GDM 56

Db 147 KQLQKAVERGAVPLEMLEIALNLNLQKQLLSEEEKNLSASTVLTAEQSFSGSLGHL 206

QY 57 PSRRNVACVCRSWRLITKEIVAPFSSKLTFPISL-----KQSGP----- 100

Db 207 QNRNNSCDSRGQTTKARFLYGE--NMPPQDAEIGYRNSLRQNTKTKQSCFGRVVPVNLN 264

QY 101 -----RDSLVOCFIKRN-----RNTOSYHLYLGLTTSL 128

Db 265 SPDCDKTDDSVVPCFMKQTSRSCRDLVPGSKPSGNDSCELNLSK----- 313

QY 129 TDNGKF---LLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFIKGTFTVFDG 185

Db 314 VQNSHFKEPLVSDKSSSELLITD-SITLK---NKTSSLLAKLEET---KEYQPEVEPE 365

QY 186 TCAAKMKQSRSSNFIKVSFVPGQSGYPIAHISYELNVLSRGRPRMRCIMDTIPMSIVS 245

Db 366 SNQKQWQSKRSECCINQNPAASSNHWQIPELARKVN-----TQ 404

QY 246 RGVWASTSISFSSRSPPVFRSHSKPLRSNSASCSDSGNNLGD-----PPLVLSNKAPR 299

Db 405 KHTTEQPVFSVSKQSPFI--STSKWFDPKSICKTPSSNTLDDYMSCFPTPVVKNDPFA 462

QY 300 W-----HEQLRCWCLNFTGRVTVASVKNPQLVAVSDC-EAGQTSR---IILOFGKVG-- 348

Db 463 CQLSTPYGPACFQOQOH-QILATPLQLNLQVLASSANECISVKGRIYSILKQIGSGSS 521

QY 349 -----KDMFTWDYGPISAFQAFALCISSEFETRIA 378

Db 522 KVFQVLNEKKQIYAIKY---VNLEADNQTLDSYRNEIA 557

RESULT 14

US-10-821-234-1635

; Sequence 1635, Application US/10821234

; Publication No. US20050255114A1

; GENERAL INFORMATION:

; APPLICANT: Labat, Ivan

; APPLICANT: Stache-Crain, Birgit

; APPLICANT: Andarmani, Susan

; APPLICANT: Tang, Y. Tom

; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia

; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234


```
; CURRENT FILING DATE: 2004-04-07
; PRIOR FILING DATE: 2004-04-07
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1635
; LENGTH: 1897
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1635

Query Match      4.0%; Score 79; DB 6; Length 1897;
Best Local Similarity 23.8%; Pred. No. 20;
Matches 85; Conservative 38; Mismatches 138; Indels 96; Gaps 19;

QY 14 PHRVHAAASTANSSDPFSW---SELPEELL-----REILIRVETVGGDWPSRRNVV 63
DB 811 PGRPTWMIISTAMNTALLQWHPKELPGELLYGRLQYCRADARPNTIDFGKDDQHFTVT 870
QY 64 AC-AGVCRWRILITKEIVAV-PEFSSKLTFFISLKQSGPRDSLVCQFIKRNRTQSYHLY 121
DB 871 GLHKGTTYIFRLAAKNRAGLGEFEKEIRTPEDLPSGFP-----QNLHV- 914
QY 122 LGTTTSLTD-----NGKFL-----AASKLKRATC-----TDY 149
DB 915 TGLTTTTELAWDPPVLAERNGRIISYTVVFRDINSQQELQNTTDTTRFTLTGLKPDTTY 974
QY 150 IISLRDDDISKRSNAYLGRMRSNFLTGTFTVPGSQTGAQKQKRSNFIKVSPPVQ 209
DB 975 DIKVRW-WTSKSGPLSPSIQSRTPVE-QVFAKNFRVAAAMKTS-----VLLSWEVPD- 1026
QY 210 SYPIA---HISYELNVLGSRGPRRCIMDTIPMS-----IVESRGVVASTSISFSSRS 262
DB 1027 SYKSAVPPFKILYNGQSVEDVGHSMRKLIALDLQNTTEYSFVLMNRGSSAGGLQHLVIRTA 1086
QY 263 PVFRSHSKPLRNSASCSGNNLG-----DPPLVL-----SNKAPRW 300
DB 1087 PDLPLH-KPLPA-SAYIEDGRFDSMPHVQDPSLVRWFVIVVPIRDRVGSMLTPRW 1141

RESULT 15
US-11-000-463-250
; Sequence 250, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785GIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870

; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 1907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-250

Query Match      4.0%; Score 79; DB 7; Length 1907;
Best Local Similarity 23.8%; Pred. No. 20;
Matches 85; Conservative 38; Mismatches 138; Indels 96; Gaps 19;

QY 14 PHRVHAAASTANSSDPFSW---SELPEELL-----REILIRVETVGGDWPSRRNVV 63
DB 821 PGRPTWMIISTAMNTALLQWHPKELPGELLYGRLQYCRADARPNTIDFGKDDQHFTVT 880
QY 64 AC-AGVCRWRILITKEIVAV-PEFSSKLTFFISLKQSGPRDSLVCQFIKRNRTQSYHLY 121
DB 881 GLHKGTTYIFRLAAKNRAGLGEFEKEIRTPEDLPSGFP-----QNLHV- 924
QY 122 LGTTTSLTD-----NGKFL-----AASKLKRATC-----TDY 149
DB 925 TGLTTTTELAWDPPVLAERNGRIISYTVVFRDINSQQELQNTTDTTRFTLTGLKPDTTY 984
QY 150 IISLRDDDISKRSNAYLGRMRSNFLTGTFTVPGSQTGAQKQKRSNFIKVSPPVQ 209
DB 985 DIKVRW-WTSKSGPLSPSIQSRTPVE-QVFAKNFRVAAAMKTS-----VLLSWEVPD- 1036
QY 210 SYPIA---HISYELNVLGSRGPRRCIMDTIPMS-----DPPLVL-----SNKAPRW 300
DB 1037 SYKSAVPPFKILYNGQSVEDVGHSMRKLIALDLQNTTEYSFVLMNRGSSAGGLQHLVIRTA 1096
QY 263 PVFRSHSKPLRNSASCSGNNLG-----DPPLVL-----SNKAPRW 300
DB 1097 PDLPLH-KPLPA-SAYIEDGRFDSMPHVQDPSLVRWFVIVVPIRDRVGSMLTPRW 1151

Search completed: December 23, 2005, 23:29:44
Job time : 17 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 23, 2005, 22:46:55 ; Search time 198 Seconds
(without alignments)
843.252 Million cell updates/sec

Title: US-10-763-042-9
Perfect score: 1962
Sequence: 1 MTFRLLQEMSRPHRVHA.....SAFQAFACLSSFETRIACE 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A: Geneseq 21: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *
9: Geneseqp2005s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1962	100.0	380	8	Adn73049 Thale cre
2	1962	100.0	380	9	Adw64863 Arabidops
3	1442.5	73.5	389	8	Ado61995 Transcrip
4	1437.5	73.3	389	3	Aag15637 Arabidops
5	1414.5	72.1	380	9	Adw64865 Arabidops
6	1409.5	71.8	380	3	Aag15638 Arabidops
7	1215.5	62.0	403	9	Adw17479 Arabidops
8	1153.5	58.8	408	9	Adw17480 Eucalyptu
9	1149	58.6	406	8	Adn73405 Thale cre
10	1149	58.6	406	9	Adw64857 Arabidops
11	1142.5	58.2	308	3	Aag15639 Arabidops
12	1119.5	57.1	411	9	Adw17849 Arabidops
13	1082	55.1	418	9	Adw17851 Arabidops
14	1049.5	53.5	418	9	Adw17850 Arabidops
15	1028	52.4	414	9	Adw18506 Arabidops
16	1028	52.4	414	9	Adw18507 Arabidops
17	1018	51.9	436	8	Ady24685 Arabidops
18	1018	51.9	436	8	Adx97136 Arabidops
19	1013.5	51.7	455	9	Adw64907 Arabidops
20	1010.5	51.5	437	8	Adx94433 Arabidops
21	1006	51.3	442	8	Adx97135 Arabidops
22	1002	51.1	412	9	Adw17848 Arabidops
23	999.5	50.9	445	7	Adx37191 Arabidops
24	999.5	50.9	445	8	Adi44433 Arabidops

ALIGNMENTS

RESULT 1	ADN73049	ADN73049 standard; protein; 380 AA.	445	9	ADW64864	ADW64864 Arabidops
ID	ADN73049	standard; protein; 380 AA.	450	9	ADW64855	ADW64855 Arabidops
XX	ADN73049		468	8	ADY06799	ADY06799 Plant ful
AC	ADN73049		424	9	ADW17484	ADW17484 Eucalyptu
XX	ADN73049		424	9	ADW17483	ADW17483 Eucalyptu
DT	15-JUL-2004	(first entry)	424	9	ADW17485	ADW17485 Eucalyptu
XX	Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 944.		415	9	ADW17481	ADW17481 Eucalyptu
DE	Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 944.		408	9	ADW17847	ADW17847 Pinus rad
XX	plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;		527	8	ADX97137	ADX97137 Plant ful
KW	animal feed product; thale cress; cell wall biosynthesis;		551	8	ADY09712	ADY09712 Plant ful
KW	nitrogen metabolism; carbon metabolism.		394	8	ADW64856	ADW64856 Arabidops
XX	Arabidopsis thaliana.		480	8	ADW48240	ADW48240 Polytypei
XX	WO2004035798-A2.		429	8	ADO61991	ADO61991 Transcrip
PN	WO2004035798-A2.		429	9	ADW64859	ADW64859 Arabidops
XX	29-APR-2004.		388	9	ADW64860	ADW64860 Arabidops
PD	29-APR-2004.		421	9	ADW17482	ADW17482 Eucalyptu
XX	20-OCT-2003; 2003WO-EP011658.		379	3	AAG07737	AAG07737 Arabidops
XX	18-OCT-2002; 2002EP-00079408.		379	9	ADW64861	ADW64861 Arabidops
PR	(CROP-) CROPDESIGN NV.		400	3	AAG07736	AAG07736 Arabidops
XX	Inze D, De Veylder L, Vlieghe K;		370	8	ADT58092	ADT58092 Plant pol
XX	WPI; 2004-348466/32.		342	3	AAG07738	AAG07738 Arabidops
DR	N-PSDB; ADN73048.					
XX	Altering plant characteristics, useful for producing plants for enzyme or					
PT	pharmaceutical production comprises modifying in a plant, expression of					
PT	one or more nucleic acids and/or modifying level or activity of one or					
PT	more proteins.					
XX	Claim 1; SEQ ID NO 944; 134pp; English.					
XX	This invention relates to a novel method for altering one or more plant					
CC	characteristics. Specifically, it refers to identifying genes that are up					
CC	-or down-regulated in transgenic plants overexpressing the heterodimeric					
CC	E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to					
CC	alter plant characteristics accordingly. The present invention describes					
CC	generating transgenic plants for the production of growth regulators,					
CC	enzymes, therapeutics, pharmaceuticals and animal feed products, where					
CC	the altered plant characteristics are selected from increased yield or					
CC	biomass, enhanced survival capacity, stress tolerance, plant architecture					

CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polypeptide sequence is thale cress protein
CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing
CC the E2Fa/Dra transcription factor, given in an exemplification of the
CC invention.
XX
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 1962; DB 8; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.4e-211;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTFRSLLOEMRSRPHRVVHAAASTANSSDPFWSSELPEELLREILLIRVETVDDGWPSPRR 60
DB 1 MTFRSLLOEMRSRPHRVVHAAASTANSSDPFWSSELPEELLREILLIRVETVDDGWPSPRR 60
QY 61 NVVACAGVCRSWRLTKEIVAVPEFSSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 120
DB 61 NVVACAGVCRSWRLTKEIVAVPEFSSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 120
QY 121 YLGTTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV 180
DB 121 YLGTTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV 180
QY 181 FDGSOTGAAMQKRSNNFKVSPVPOGSPYPIAHISYELNVLGSRGPRRMCIMDTIPM 240
DB 181 FDGSOTGAAMQKRSNNFKVSPVPOGSPYPIAHISYELNVLGSRGPRRMCIMDTIPM 240
QY 241 SIVESRGVAVASTSISSFSSRSPPVFRSHSKPLRSNSASCSDSGNNLGDPLVLNKAAPRW 300
DB 241 SIVESRGVAVASTSISSFSSRSPPVFRSHSKPLRSNSASCSDSGNNLGDPLVLNKAAPRW 300
QY 301 HEQLRCWCLNFHGRVTVASVKNFQLVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
DB 301 HEQLRCWCLNFHGRVTVASVKNFQLVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
QY 361 SAFOAFAICLSSPETRIACE 380
DB 361 SAFOAFAICLSSPETRIACE 380

RESULT 2
ADM64863
ID ADM64863 standard; protein; 380 AA.

XX ADM64863;

XX 07-APR-2005 (first entry)

XX Arabidopsis thaliana TUBBY-like protein (TLP) 9.

XX Transgenic plant; salt tolerance; crop improvement; cold tolerance;
KW oxidative stress; drought resistance; TUBBY-like protein; plant.

XX Arabidopsis thaliana.

XX US2005014266-A1.

XX 20-JAN-2005.

XX 21-JAN-2004; 2004US-00763042.

XX 21-JAN-2003; 2003US-0441380P.

XX (SINT-) ACAD SINICA.

XX Shaw J, Lai C;

XX WPI; 2005-080955/09.
DR N-PSDB; ADM64874.
XX New isolated Arabidopsis TUBBY-like proteins, useful for producing to
PT transgenic or transformed cells or plants having higher tolerance to
PT salt, chilling, pathogens, oxidative stress, or water-deficit.
XX
PS Claim 1; SEQ ID NO 9; 43pp; English.
XX
CC The invention relates to Arabidopsis TUBBY-like proteins (TLP) and their
CC corresponding nucleic acid sequences. The invention also relates to a
CC transformed cell or a transgenic plant containing TLP nucleic acid and a
CC transformed cell or a transgenic plant lacking TLP nucleic acid. The TLP
CC polypeptides and nucleic acids are useful for producing transgenic or
CC transformed cells or plants having higher tolerance to salt, chilling, the
CC pathogens, oxidative stress or water-deficit. The present sequence is the
CC Arabidopsis thaliana TLP 9 protein.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 1962; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.4e-211;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTFRSLLOEMRSRPHRVVHAAASTANSSDPFWSSELPEELLREILLIRVETVDDGWPSPRR 60
DB 1 MTFRSLLOEMRSRPHRVVHAAASTANSSDPFWSSELPEELLREILLIRVETVDDGWPSPRR 60
QY 61 NVVACAGVCRSWRLTKEIVAVPEFSSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 120
DB 61 NVVACAGVCRSWRLTKEIVAVPEFSSKLTFFPISLKQSGPRDSLVOCFIKRNRNTQSYHL 120
QY 121 YLGTTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV 180
DB 121 YLGTTSLTDNGKFLAASKLKRATCTDYIISLRSDDISKRSNAYLGRMRSNFLTGTFTV 180
QY 181 FDGSOTGAAMQKRSNNFKVSPVPOGSPYPIAHISYELNVLGSRGPRRMCIMDTIPM 240
DB 181 FDGSOTGAAMQKRSNNFKVSPVPOGSPYPIAHISYELNVLGSRGPRRMCIMDTIPM 240
QY 241 SIVESRGVAVASTSISSFSSRSPPVFRSHSKPLRSNSASCSDSGNNLGDPLVLNKAAPRW 300
DB 241 SIVESRGVAVASTSISSFSSRSPPVFRSHSKPLRSNSASCSDSGNNLGDPLVLNKAAPRW 300
QY 301 HEQLRCWCLNFHGRVTVASVKNFQLVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
DB 301 HEQLRCWCLNFHGRVTVASVKNFQLVAVSDCEAGQTSERIILQFGKVGKDMFTMDYGYPI 360
QY 361 SAFOAFAICLSSPETRIACE 380
DB 361 SAFOAFAICLSSPETRIACE 380

RESULT 3
ADO61995
ID ADO61995 standard; protein; 389 AA.

XX ADO61995;

XX 15-JUL-2004 (first entry)

XX Transcription factor G2603, SEQ ID 462.

XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development.

XX Arabidopsis thaliana.

XX WO2004031349-A2.

XX

PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144633P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155133P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159291P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

PR	12-OCT-1999;	99US-0158369P.	KW	plant; transcription; gene regulation; gene expression; transgenic plant;
PR	13-OCT-1999;	99US-0159293P.	KW	drought resistance; disease resistance; salt tolerance; cold tolerance;
PR	13-OCT-1999;	99US-0159294P.	XX	freezing tolerance; flowering; flavor enhancer; flower color.
PR	13-OCT-1999;	99US-0159295P.	OS	Eucalyptus grandis.
PR	14-OCT-1999;	99US-0159329P.	XX	WO200501050-A2.
PR	14-OCT-1999;	99US-0159330P.	PN	06-JAN-2005.
PR	14-OCT-1999;	99US-0159331P.	XX	07-JUN-2004; 2004WO-US017965.
PR	14-OCT-1999;	99US-0159637P.	XX	06-JUN-2003; 2003US-0476189P.
PR	14-OCT-1999;	99US-0159638P.	PD	(ARBO-) ARBORGEN LLC.
PR	18-OCT-1999;	99US-0159584P.	XX	Blokeberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PR	21-OCT-1999;	99US-0160741P.	PI	Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PR	21-OCT-1999;	99US-0160767P.	PI	Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
PR	21-OCT-1999;	99US-0160770P.	XX	WPI; 2005-075542/08.
PR	21-OCT-1999;	99US-0160814P.	DR	N-PSDB; ADW16693.
PR	21-OCT-1999;	99US-0160815P.	XX	New polynucleotides isolated from plants encoding transcription factors,
PR	22-OCT-1999;	99US-0160980P.	PT	and polypeptides encoded by such polynucleotides, useful for regulating
PR	22-OCT-1999;	99US-0160981P.	PT	gene transcription and gene expression.
PR	22-OCT-1999;	99US-0160989P.	XX	Claim 31; SEQ ID NO 1241; 1265pp; English.
PR	25-OCT-1999;	99US-0161404P.	CC	This invention relates to novel isolated plant nucleic acid molecules, or
PR	25-OCT-1999;	99US-0161405P.	CC	variants thereof, that encode transcription factors. Specifically, it
PR	25-OCT-1999;	99US-0161406P.	CC	refers to transcription factor proteins that are capable of binding to
PR	26-OCT-1999;	99US-0161359P.	CC	DNA in order to regulate gene transcription and gene expression in a
PR	26-OCT-1999;	99US-0161360P.	CC	plants, in particular Eucalyptus grandis and Pinus radiata. The present
PR	26-OCT-1999;	99US-0161361P.	CC	invention describes DNA constructs containing DNA encoding a
PR	28-OCT-1999;	99US-0161920P.	CC	transcription factor that regulates the promoter, which is operably
PR	28-OCT-1999;	99US-0161922P.	CC	linked to the desired nucleic acid to be expressed. It further provides
PR	28-OCT-1999;	99US-0161993P.	CC	transgenic plants expressing a transcription factor that confers a trait
PR	29-OCT-1999;	99US-0162142P.	CC	to the plant such as increased drought, salt or disease tolerance, height
Query Match	71.8%;	Score 1409.5; DB 3; Length 380;	CC	change, enhanced cold/ frost tolerance, enhanced color, health and
Best Local Similarity	72.6%;	Pred. No. 7.8e-149;	CC	nutritional characteristics, as well as improved taste, starch
Matches 281; Conservative 31; Mismatches 52; Indels 23; Gaps 5;			CC	composition, flower longevity and germination, amongst others.
QY	10 MRSRPHRVH-----AAASTANSDFPFWSELPPEELLREILIRVETVDGDPWPSRRNVV	63	CC	Accordingly, such plants that are successfully transfected with a DNA
Db	1 MRSRPHRVHDLAAADSTSVSSQDYRWSEIPEELLREILIRVEADGGWPSRRSVV	60	CC	construct can be characterized by a difference in flower color, petal or
QY	64 ACAGVCRWRILTKETIVAVPEFSKLTFTPIISLKQSGPRDSLVQCFTKRNNTOSYHLYLG	123	CC	leaf shape and size, aroma or plant height. This polypeptide is a plant
Db	61 ACAGVCRWRLLMNETVWVWVPEISSKLTFTPIISLKQSGPRDSLVQCFTKRNNTOSYHLYLG	120	CC	transcription factor protein sequence of the invention.
QY	124 LITSLTNCKEFLAASKLRATCTDYIIISLRDDISKRSNAYLGRMSNPLGKTFTVFDG	183	QY	1 MTFRSLLOEMRSRPHRVH-----AAASTANSDFPFWSELPPEELLR 42
Db	121 LTNLSLTDGKFLAACKLHCTCTDYIIISLRDDMSRRSQVYGVKRSNPLGKTFTVFDG	180	Db	1 MPKSLIHMRSRRAVHDGSAADAAAAAEEAAARAAADGLSRCAWNPPELLR 60
QY	184 S---QTGAARKWQSRSNFTKVSPPVPGSQSYPIAHISYELNVLGSRGPRRMRCIMDTIPM	240	QY	43 EILIRVETVDGDPWPSRRNVVACAGVCRWRILTKETIVAVPEFSKLTFTPIISLKQSGPRD
Db	181 NLLPSTGAARKLRSYNAKYSAPVPLGSPYVAHITYELNVLGSRGPRRMRCIMDTIPT	240	Db	61 EVLVKIEASE-GSWFSKSVVACAGVCRWRHLVKEIVKVPVSGILTFPIISVKQPGPRE
QY	241 SIVSRGVVASTSISSFSRRSPVFRSHKPLRSNSASCSDSGNILGDPPLVLSNKAPRW	300	QY	103 SLVQCFIKRNNTOSYHLYLGTLTSTNGKFLAASKLRATCTDYIIISLRDDISKRS 162
Db	241 STMPCQVASEPSEFPPLGTRSTLSRSQSKPLRSS-----SHLKETPLVLSNKTFRW	293	Db	120 SLLQCFIKENRSNQTVHLYGLTNALSDDGKFLAARKCRRTCTDYVLSLHADDVSKGS
QY	301 HEQLRCWCLNFCGRVTVASVKNFOLVAV-SDCEAG-----QTSERIILQFGVKGDMFT	353	QY	163 NAYLGRMSNPLGKTFTVFDGSGQT-GAAKWQSRSNFEI---KVSPPVPGSQSYPIAHISY
Db	294 HEQLRCWCLNFCGRVTVASVKNFOLVAVAGAGSCSGTGMSPEROSRIILQFGVKGDMFT	353	Db	180 CTYVGLKRSNPLGKTFTVFDGSGQPPHFGAKMPSRSLVSSKQVSPRAPAGNPFVAHISY
QY	354 MDYGYPISAFOAFAICLSSFETRIACE 380		QY	219 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS
Db	354 MDYGYPISAFOAFAICLSSFETRIACE 380		Db	240 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS
RESULT 7			QY	240 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS
ADW17479			Db	240 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS
ID	ADW17479 standard; protein; 403 AA.		QY	240 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS
XX	ADW17479;		Db	240 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS
AC	ADW17479;		QY	240 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS
XX	24-MAR-2005 (first entry)		Db	240 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS
DT	Eucalyptus grandis transcription factor protein TUBBY family Seq 1241.		QY	240 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS
XX			Db	240 ELNVLGSRGPRRMRCIMDTIPMSIVSRGVVASTSISSFS-----RSSPFRSHSKPLRS

QY	275	NSASCDSDGNLGG--DPPLVLSNKAPRWHEQLRCWCLNFHGRVTVASVKNFOLVAVSDCE	332	Db	3	SFKSIQDVKDBIGNISRKGFHVKFGYGLASRSRSHRVVQDSSVVVDALKQSCWASMPPEL	62
Db	295	NRAESVSSPLVSQKDGALVLRNKAPRWHEQLRCWCLNFHGRVTVASVKNFOLVAVSDCE	354	QY	41	LREILIRVETDGDWPSRRNVACAGVCRSRWILTKETIVAVPEFSSKLTFFPISLKQSGP	100
QY	333	-AGQTSERIILQFGKVGKDMFTMDYGPISAFQAFACLSSPETRIACE	380	Db	63	LRDVLMRIKSE-STWPPRKNVACAGVCRNRWREIMKSVIKTPEVSGKLTFFPISLKQSGP	121
Db	355	PAGPQHEKIILOFGKVGKDLFTMDYRYPISAFQAFACLSSPETRIACE	403	QY	101	RDSLVOCFIKENRNTQSYHLYGLTSLTDNGKFLAASKLKRACTDYIISLSDDISK	160
RESULT 8				Db	122	RDSLQCTYIKNRSNQTYLYLSLQALNEDGKFLAARKLRPTCTDYVLSLNDMSK	181
ADW17480				QY	161	RSNAYLGRMRNFLTGFVFDGSGTGA-AKMOKSRSS---NFIKVSRRVPOGSGYPIAH	216
ID	ADW17480	standard; protein; 408 AA.		Db	182	GSSTYIGKLRNFLTGFVFDGSGTGA-AKMOKSRSS---NFIKVSRRVPOGSGYPIAH	241
AC	ADW17480;			QY	217	SYELNVLSGRGPRRMRCTMDIPMSIVESRGVV-----ASTSISFSRSSPFRSHSK	270
XX	24-MAR-2005	(first entry)		Db	242	SYELNVLSGRGPRRMRCTMDIPMSIVESRGVV-----ASTSISFSRSSPFRSHSK	296
DE	Eucalyptus grandis	transcription factor protein TUBBY family seq 1242.		QY	271	PLRSNSASCDSDGNLGGPP---LVLNKAAPRWHEQLRCWCLNFHGRVTVASVKNFOLVA	327
XX	plant; transcription; gene regulation; gene expression; transgenic plant;			Db	297	--KSTRVENLESGRPMSAPKQGLVLRNKAAPRWHEQLRCWCLNFHGRVTVASVKNFOLVA	354
KW	drought resistance; disease resistance; salt tolerance; cold tolerance;			QY	328	VSD-CEAGQTSERIILQFGKVGKDMFTMDYGPISAFQAFACLSSPETRIACE	380
KW	freezing tolerance; flowering; flavor enhancer; flower color.			Db	355	SPEQGVAGTEHEHNVILQFGKVGKDVFTMDYQYPISAFQSFACLSSFDTRIACE	408
XX	Eucalyptus grandis.			RESULT 9			
XX	WO2005001050-A2.			ADW173405			
PN	06-JAN-2005.			ID	ADW173405	standard; protein; 406 AA.	
XX	07-JUN-2004; 2004WO-US017965.			XX	AC	ADW173405;	
XX	06-JUN-2003; 2003US-0476189P.			XX	DT	15-JUL-2004	(first entry)
PA	(ARBO-) ARBORGEN LLC.			XX	DE	Thale cress protein upregulated in E2Fa/Dpa expressing plants seqID 1300.	
XX	Blokeberg LM, Bryant C, Connett MB, Emerson SJ, Frost MJ;			XX	KW	plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;	
PI	Forster RLS, Grigor M, Higgins C, Laeham A, Lund ST, Magnus A;			KW	KW	animal feed product; thale cress; cell wall biosynthesis;	
PI	Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;			XX	KW	nitrogen metabolism; carbon metabolism.	
DR	WPI; 2005-075542/08.			OS	OS	Arabidopsis thaliana.	
DR	N-PSDB; ADW16694.			XX	PN	WO2004035798-A2.	
XX	New polynucleotides isolated from plants encoding transcription factors,			XX	PD	29-APR-2004.	
PT	and polypeptides encoded by such polynucleotides, useful for regulating			XX	PF	20-OCT-2003; 2003WO-EP011658.	
PT	gene transcription and gene expression.			XX	PR	18-OCT-2002; 2002EP-00079408.	
XX	Claim 31; SEQ ID NO 1242; 1265pp; English.			XX	PA	(CROP-) CROPDESIGN NV.	
XX	This invention relates to novel isolated plant nucleic acid molecules, or			XX	PI	Inze D, De Veylder L, Vlieghe K;	
CC	variants thereof, that encode transcription factors. Specifically, it			XX	XX	WPI; 2004-348466/32.	
CC	refers to transcription factor proteins that are capable of binding to			DR	DR	N-PSDB; ADW173404.	
CC	DNA in order to regulate gene transcription and gene expression in a			XX	XX	Altering plant characteristics, useful for producing plants for enzyme or	
CC	plants, in particular Eucalyptus grandis and Pinus radiata. The present			XX	PT	pharmaceutical production comprises modifying in a plant, expression of	
CC	invention describes DNA constructs containing DNA encoding a			XX	PT	one or more nucleic acids and/or modifying level or activity of one or	
CC	transcription factor that regulates the promoter, which is operably			XX	PT	more proteins.	
CC	linked to the desired nucleic acid to be expressed. It further provides			XX	XX	Claim 1; SEQ ID NO 1300; 134pp; English.	
CC	transgenic plants expressing a transcription factor that confers a trait			XX	XX	This invention relates to a novel method for altering one or more plant	
CC	to the plant such as increased drought, salt or disease tolerance, height			XX	XX	characteristics. Specifically, it refers to identifying genes that are up	
CC	change, enhanced cold/frost tolerance, enhanced color, health and			XX	XX	-or down-regulated in transgenic plants overexpressing the heterodimeric	
CC	nutritional characteristics, as well as improved taste, starch			XX	XX	E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to	
CC	composition, flower longevity and germination, amongst others.			XX	XX	alter plant characteristics accordingly. The present invention describes	
CC	Accordingly, such plants that are successfully transfected with a DNA			XX	XX	generating transgenic plants for the production of growth regulators,	
CC	construct can be characterized by a difference in flower color, petal or			XX	XX	enzymes, therapeutics, pharmaceuticals and animal feed products, where	
CC	leaf shape and size, aroma or plant height. This polypeptide is a plant			XX	XX		
CC	transcription factor protein sequence of the invention.			XX	XX		
XX	Sequence 408 AA;			XX	XX		
XX	Query Match	58.8%; Score 1153.5; DB 9; Length 408;		XX	XX		
XX	Best Local Similarity	56.0%; Pred. No. 5.5e-120;		XX	XX		
XX	Matches 233; Conservative	60; Mismatches 79; Indels 43; Gaps 9;		XX	XX		
XX	2 TFRSLQGE-----MRSRPHRVVHAASSTANSSDPFSNSLPEEL 40			XX	XX		

CC the altered plant characteristics are selected from increased yield or
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture
 CC or physiology, altered endoreplication, biochemistry, signal
 CC transduction, storage lipid mobilization and/or altered photosynthesis,
 CC each relative to the corresponding wild type plants. Accordingly, these
 CC sequences can also be useful as positive or negative selectable markers
 CC during transformation of cells or tissues. The identified genes play a
 CC role in a variety of biological processes such as DNA replication, cell
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
 CC transcription factors. This polypeptide sequence is thale cress protein
 CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing
 CC the E2Fa/Dpa transcription factor, given in an exemplification of the
 CC invention.

XX Sequence 406 AA;

Query Match 58.6%; Score 1149; DB 8; Length 406;
 Best Local Similarity 56.7%; Pred. No. 1.8e-119;
 Matches 233; Conservative 58; Mismatches 84; Indels 36; Gaps 9;

QY 1 MTFRSLLOEM-----RSRPHRVVHAAASTANGSDPFSWSELPEELL 41
 DB 1 MSFKSLIQDMRGELGSLSRKGFDFVRFGYGRSRQRVVQDTSVPDAFKQSCWASMPPELL 60
 QY 42 REILIRVETVDCGDWPSRENVACVCSWRILTKETIVAVPEFSSKLTFFPISLKQSGPR 101
 DB 61 RDVLMRIEQSE-DTWPSRKNVSCAGVCNRWREIVKEIVRVPSELKLTFFPISLKQSGPR 119
 QY 102 DSLVQCFIKRNTQSYHLGLTTSLT-DNGKFLAASKLKRATCTDYIISLRSDDISK 160
 DB 120 GSLVQCYIMRNSNTQYIYLGINQAASNDGKFLAAKFRPTCTDYIISLNCDDVSR 179
 QY 161 RSNAYLGRNRSNLTGFTKFTVFGSGQTG-AAKQKSRSSNFI---KVSRRVPOGSSYPIAH 216
 DB 180 GSNTRYIGKLSNLTGFTKFTVYDAQTPNPGTVTRTRSSRLLSLKQVSRIPSGNYVAHI 239
 QY 217 SYELNVLSGRGPRRRCINDTIPMSIVESRGV-----VASTSISFSSSRSPVFRSHSK 270
 DB 240 SYELNVLSGRGPRRRCINDTIPMSIVESRGV-----VASTSISFSSSRSPVFRSHSK 295
 QY 271 PLRSNSASCSDSGNNLGDPLVLSNKA PRWHEQLRCWCLNFHGRVTVASVKNFQLVAV-S 329
 DB 296 STRASLSPGSPSAQAQEGLLVLKKA PRWHEQLRCWCLNFHGRVTVASVKNFQLVAV-S 355
 QY 330 DCEAGQTSERIILQFGKVGKDMFTMDYGYPISAFQAFALCLSFETRIACE 380
 DB 356 NGPAGPEHENVILQFGKVGKDMFTMDYGYPISAFQAFALCLSFETRIACE 406

RESULT 10

ADW64857
 ID ADW64857 standard; protein; 406 AA.

AC ADW64857;

DT 07-APR-2005 (first entry)

DE Arabidopsis thaliana TUBBY-like protein (TLP) 3.

KW Transgenic plant; salt tolerance; crop improvement; cold tolerance;
 KW oxidative stress; drought resistance; TUBBY-like protein; plant.

XX Arabidopsis thaliana.

XX US2005014266-A1.

XX 20-JAN-2005.

XX 21-JAN-2004; 2004US-00763042.

XX 21-JAN-2003; 2003US-0441380P.

PA (SINI-) ACAD SINICA.

XX

PI Shaw J, Lai C;

XX WPI; 2005-080955/09.

DR N-PSDB; ADW64868.

XX New isolated Arabidopsis TUBBY-like proteins, useful for producing
 PT transgenic or transformed cells or plants having higher tolerance to
 PT salt, chilling, pathogens, oxidative stress, or water-deficit.

PS Claim 1; SEQ ID NO 3; 43pp; English.

XX The invention relates to Arabidopsis TUBBY-like proteins (TLP) and their
 CC corresponding nucleic acid sequences. The invention also relates to a
 CC transformed cell or a transgenic plant containing TLP nucleic acid and a
 CC transformed cell or a transgenic plant lacking TLP nucleic acid. The TLP
 CC polypeptides and nucleic acids are useful for producing transgenic or
 CC transformed cells or plants having higher tolerance to salt, chilling,
 CC pathogens, oxidative stress or water-deficit. The present sequence is the
 CC Arabidopsis thaliana TLP 3 protein.

XX Sequence 406 AA;

Query Match 58.6%; Score 1149; DB 9; Length 406;
 Best Local Similarity 56.7%; Pred. No. 1.8e-119;
 Matches 233; Conservative 58; Mismatches 84; Indels 36; Gaps 9;

QY 1 MTFRSLLOEM-----RSRPHRVVHAAASTANGSDPFSWSELPEELL 41

DB 1 MSFKSLIQDMRGELGSLSRKGFDFVRFGYGRSRQRVVQDTSVPDAFKQSCWASMPPELL 60

QY 42 REILIRVETVDCGDWPSRENVACVCSWRILTKETIVAVPEFSSKLTFFPISLKQSGPR 101

DB 61 RDVLMRIEQSE-DTWPSRKNVSCAGVCNRWREIVKEIVRVPSELKLTFFPISLKQSGPR 119

QY 102 DSLVQCFIKRNTQSYHLGLTTSLT-DNGKFLAASKLKRATCTDYIISLRSDDISK 160

DB 120 GSLVQCYIMRNSNTQYIYLGINQAASNDGKFLAAKFRPTCTDYIISLNCDDVSR 179

QY 161 RSNAYLGRNRSNLTGFTKFTVFGSGQTG-AAKQKSRSSNFI---KVSRRVPOGSSYPIAH 216

DB 180 GSNTRYIGKLSNLTGFTKFTVYDAQTPNPGTVTRTRSSRLLSLKQVSRIPSGNYVAHI 239

QY 217 SYELNVLSGRGPRRRCINDTIPMSIVESRGV-----VASTSISFSSSRSPVFRSHSK 270

DB 240 SYELNVLSGRGPRRRCINDTIPMSIVESRGV-----VASTSISFSSSRSPVFRSHSK 295

QY 271 PLRSNSASCSDSGNNLGDPLVLSNKA PRWHEQLRCWCLNFHGRVTVASVKNFQLVAV-S 329

DB 296 STRASLSPGSPSAQAQEGLLVLKKA PRWHEQLRCWCLNFHGRVTVASVKNFQLVAV-S 355

QY 330 DCEAGQTSERIILQFGKVGKDMFTMDYGYPISAFQAFALCLSFETRIACE 380

DB 356 NGPAGPEHENVILQFGKVGKDMFTMDYGYPISAFQAFALCLSFETRIACE 406

RESULT 11

AAGI5639

ID AAGI5639 standard; protein; 308 AA.

XX AAGI5639;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 15970.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX	06-SEP-2000.	PR	02-JUL-1999;	99US-0142055P.
PD		PR	06-JUL-1999;	99US-0142390P.
XX		PR	08-JUL-1999;	99US-0142803P.
PF	25-FEB-2000;	PR	09-JUL-1999;	99US-0142920P.
XX		PR	12-JUL-1999;	99US-0142977P.
PR	25-FEB-1999;	PR	13-JUL-1999;	99US-0143542P.
PR	05-MAR-1999;	PR	14-JUL-1999;	99US-0144005P.
PR	09-MAR-1999;	PR	15-JUL-1999;	99US-0144085P.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086P.
PR	23-MAR-1999;	PR	19-JUL-1999;	99US-0144325P.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331P.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332P.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333P.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334P.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335P.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352P.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884P.
PR	23-APR-1999;	PR	21-JUL-1999;	99US-0144814P.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088P.
PR	30-APR-1999;	PR	22-JUL-1999;	99US-0145085P.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145087P.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192P.
PR	06-MAY-1999;	PR	23-JUL-1999;	99US-0145145P.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145218P.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	PR	11-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	PR	12-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341P.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148565P.
PR	18-JUN-1999;	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066P.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	PR		

```
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 58.2%; Score 1142.5; DB 3; Length 308;
Best Local Similarity 73.2%; Pred. No. 6e-119;
Matches 229; Conservative 26; Mismatches 41; Indels 17; Gaps 4;

QY 78 EIVAVPEFSKLTFTPSLKQSGPRDSLVCQFTIKRNRNTQSYHLYLGLTSLTDNGKFLLA 137
Db 3 ETWVPEISSKLTFTPSLKQSGPRDSLVCQFTIKRNRITQSYHLYLGLTSLTDNGKFLLA 62
QY 138 ASKLRATCTDYIISLRSDDISKRSNAYLGRMSNFLTCTKFTVFDGS---QTGAAMQKS 194
Db 63 ACKLHTTCTDYIISLRSDDISKRSNAYLGRMSNFLTCTKFTVFDGNLLPSTGAAMLRKS 122
QY 195 RSSNFTKSPRVPGQSYPTAHISYELNVLGSRGPRMRCIMDTIPMSIVESRGVASTSI 254
Db 123 RSNPAKVAKVPLGSPYVAHIYELNVLGSRGPRMRCIMDTIPSTWEPQVASEPSE 182
QY 255 SSFSSRSSVFRSHSKPLRSNSASCSDSGNNLGDPEPLVLSNKAAPRWHEQLRCWCLNFHGR 314
Db 183 FPLIGTRSTLSRSQSKPLRSS-----SHLKETPLVLSNKAAPRWHEQLRCWCLNFHGR 235
QY 315 VTVASVKNFQVAV-SDCAG-----QTSERILOFGKVGKMDFTMDYGYPISAQFA 367
Db 236 VTVASVKNFQVAAAGASCSGTGMSPEROSERILOFGKVGKMDFTMDYGYPISAQFA 295
QY 368 ICLSSPETRIACE 380
Db 296 ICLSSPETRIACE 308

RESULT 12
ADM17849
ID ADM17849 standard; protein; 411 AA.
XX
```

```
AC ADM17849;
XX 24-MAR-2005 (first entry)
XX DE Pinus radiata transcription factor protein TUBBY family Seq 1627.
XX KW plant; transcription; gene regulation; gene expression; transgenic plant;
XX KW drought resistance; disease resistance; salt tolerance; cold tolerance;
XX KW freezing tolerance; flowering; flavor enhancer; flower color.
OS Pinus radiata.
XX PN WO2005001050-A2.
XX PD 06-JAN-2005.
XX PF 07-JUN-2004; 2004WO-US017965.
XX PR 06-JUN-2003; 2003US-0476189P.
XX PA (ARBO-) ARBorgen LLC.
XX PI Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX WPI; 2005-075542/08.
DR N-PSDB; ADM17063.
XX New polynucleotides isolated from plants encoding transcription factors,
XX and polypeptides encoded by such polynucleotides, useful for regulating
XX gene transcription and gene expression.
PS Claim 31; SEQ ID NO 1627; 1265pp; English.
CC This invention relates to novel isolated plant nucleic acid molecules, or
CC variants thereof, that encode transcription factors. Specifically, it
CC refers to transcription factor proteins that are capable of binding to
CC DNA in order to regulate gene transcription and gene expression in a
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC invention describes DNA constructs containing DNA encoding a
CC transcription factor that regulates the promoter, which is operably
CC linked to the desired nucleic acid to be expressed. It further provides
CC transgenic plants expressing a transcription factor that confers a trait
CC to the plant such as increased drought, salt or disease tolerance, height
CC change, enhanced cold/frost tolerance, enhanced color, health and
CC nutritional characteristics, as well as improved taste, starch
CC composition, flower longevity and germination, amongst others.
CC Accordingly, such plants that are successfully transfected with a DNA
CC construct can be characterized by a difference in flower color, petal or
CC leaf shape and size, aroma or plant height. This polypeptide is a plant
CC transcription factor protein sequence of the invention.
SQ Sequence 411 AA;
Query Match 57.1%; Score 1119.5; DB 9; Length 411;
Best Local Similarity 54.4%; Pred. No. 3.7e-116;
Matches 224; Conservative 65; Mismatches 90; Indels 33; Gaps 7;
QY 1 MTFRSLQEQ-----MRSPRHRVHAAASTANSSDPFS-WSLPEE 39
Db 1 MGFRLTLQDVRGIGISIRSRSLDVKFSYLSRSHSAVHDSHYRSQVSLPQSCWANNPE 60
QY 40 LLREILIRVETVDGDPSPRRNVNACAGVCRSWRLITKEIVAVPEFSKLTFTPSLKQSG 99
Db 61 LLRDVISRLEASEGA-WPARKYVACSGVCRSWRAITKEIVQTPQLSGKLTFTPSLRQPG 119
QY 100 PRDSLVCQFTIKRNRNTQSYHLYLGLTSLTDNGKFLLAASKLRATCTDYIISLRSDDIS 159
Db 120 PRDFLTQCFIKRDRATSSYHLYLGLTSLTDNGKFLLAARKSRRTCTDYIISLRDAMS 179
QY 160 KESNAYLGRMSNFLTCTKFTVFDGSQ--TG--AAKMQKSRSSNFIVKSPRVPGQSYPIAH 215
```

180 RGSNEYIGKLRNGLTKFTIYDSQPCGTGATSTSGRRVSKQVSPRVAGSVVAH 239
216 ISEYELNVLGSRGPRMRRCIMDTIPMSIVESRGVAVASTSISFSSRSPPFRSH----SKP 271
240 IAYELNVLGTRGPRMRQCILDVAPASAMEAGGSAPTTECPPLSSLDLSTLPHLGLKSEV 299
272 LRSNSASCSDSGNLGDPLVLNKNAPRWHEQLRCWCLNFHGRVTVASVKNFQVAVSD- 330
300 IETSPFGPIINSNWKREGPLLLKNAPRWHEQLQCWCLNFKGRVTVPSVKNFQVAAATEP 359
331 --CEAGQTSERIILQFGVKGMFTMDYGPISAFQAFACILSSPETRIACE 380
360 TQPAQSDHDKVILQFGVKGMFTMDYRYPLSAFOAFACILSSFDTKLACE 411

RESULT 13
ADW17851
ID ADW17851 standard; protein; 418 AA.
XX
AC ADW17851;
XX
XX
DT 24-MAR-2005 (first entry)
DE pinus radiata transcription factor protein TUBBY family Seq 1629.
XX
XX plant; transcription; gene regulation; gene expression; transgenic plant;
KW drought resistance; disease resistance; salt tolerance; cold tolerance;
KW freezing tolerance; flowering; flavor enhancer; flower color.
XX
XX Pinus radiata.
XX
XX WO2005001050-A2.
XX
XX 06-JAN-2005.
XX
XX 07-JUN-2004; 2004WO-US017965.
XX
XX 06-JUN-2003; 2003US-0476189P.
XX
XX (ARBO-) ARBORGEN LLC.
XX
XX Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX
XX WPI; 2005-075542/08.
XX
XX N-PSDB; ADW17065.

XX
XX New polynucleotides isolated from plants encoding transcription factors,
PT and polypeptides encoded by such polynucleotides, useful for regulating
PT gene transcription and gene expression.
XX
XX Claim 31; SEQ ID NO 1629; 1265pp; English.
XX
XX This invention relates to novel isolated plant nucleic acid molecules, or
CC variants thereof, that encode transcription factors. Specifically, it
CC refers to transcription factor proteins that are capable of binding to
CC DNA in order to regulate gene transcription and gene expression in a
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC invention describes DNA constructs containing DNA encoding a
CC transcription factor that regulates the promoter, which is operably
CC linked to the desired nucleic acid to be expressed. It further provides
CC transgenic plants expressing a transcription factor that confers a trait
CC to the plant such as increased drought, salt or disease tolerance, height
CC change, enhanced cold/frost tolerance, enhanced color, health and
CC nutritional characteristics, as well as improved taste, starch
CC composition, flower longevity and germination, amongst others.
CC Accordingly, such plants that are successfully transfected with a DNA
CC construct can be characterized by a difference in flower color, petal or
CC leaf shape and size, aroma or plant height. This polypeptide is a plant
CC transcription factor protein sequence of the invention.
XX
XX Sequence 418 AA;

Query Match 55.1%; Score 1082; DB 9; Length 418;
Best Local Similarity 53.2%; Pred. No. 6.3e-112; Indels 52; Gaps 8;
Matches 226; Conservative 59; Mismatches 88;
QY 1 MTFRSLQLQEM-----RSRRPHRVV-----HAAASTANSDDPFS-----W 33
DB 1 MTVXNIFQDMKSGISISRRSFEVKFLRSQSQAVDDLHYRSNANNKNNNGGLQCSILW 60
QY 34 SELPEELREILIRVETVDGDNWPSRRNVACAGVCRSWRLITKEIVAVPFPSSKLTPTPI 93
DB 61 ANMPELIRDIVQRIESE-SSWPFKSVVACAGVCMTRQITKELVPIPELSGKLTPTPI 119
QY 94 SLKQSGPDSLVQCFIKENRNTQSHLYLGLTTSITDNGKFLLAASKLKRACTDYIISL 153
DB 120 SLKOPGSPSPVQCFIKDRATSTYQLFCLNTTILSENGKFLLAARKFRRTATSTYIISL 179
QY 154 RSDDISKSNAYLGRMRSNFLTGTFTVPDG-----SQTGAAMQKRSRNSFIKVSPPVPG 209
DB 180 NADDMRGSTNYVGLKRSNFLTGTFTYDPSPPHSGAIASTRAGRRVYGAKQVSPRPAG 239
QY 210 SYPIAHISYELNVLGSRGPRMRRCIMDTIPMSIVESRG-----VVASTSISFSSRSPPV 264
DB 240 SYNTIAHIGYELNVLGTRGPRMRQCIMHSIPASAVQPGGSAPTPVPSPVASLDQSLASCT 299
QY 265 FRSHKPLRSNSASCSDSGNLGDPLVLNKNAPRWHEQLRCWCLNFHGRVTVASVKNFQ 324
DB 300 FGSKSCILSSSISSEPLTGSNLKDEPLVLNKNAPRWHEQLQCWCLNFKGRVTVASVKNFQ 359
QY 325 LVAV-----SDCEAGQTSERIILQFGVKGMFTMDYGPISAFQAFACILSSPET 375
DB 360 LVAPEPNPSVQNDK-----DKVILQFGVKGMFTMDYRYPLSAFOAFACILSSPET 413
QY 376 RIACE 380
DB 414 KLACE 418
RESULT 14
ADW17850
ID ADW17850 standard; protein; 418 AA.
XX
AC ADW17850;
XX
DT 24-MAR-2005 (first entry)
XX
DE Pinus radiata transcription factor protein TUBBY family Seq 1628.
XX
XX plant; transcription; gene regulation; gene expression; transgenic plant;
KW drought resistance; disease resistance; salt tolerance; cold tolerance;
KW freezing tolerance; flowering; flavor enhancer; flower color.
XX
XX Pinus radiata.
XX
XX WO2005001050-A2.
XX
XX 06-JAN-2005.
XX
XX 07-JUN-2004; 2004WO-US017965.
XX
XX 06-JUN-2003; 2003US-0476189P.
XX
XX (ARBO-) ARBORGEN LLC.
XX
XX Bloksberg LN, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX
XX WPI; 2005-075542/08.
XX
XX N-PSDB; ADW17064.
XX
XX New polynucleotides isolated from plants encoding transcription factors,
PT and polypeptides encoded by such polynucleotides, useful for regulating

PT gene transcription and gene expression.
PS Claim 31; SEQ ID NO 1628; 1265pp; English.
XX
CC This invention relates to novel isolated plant nucleic acid molecules, or
CC variants thereof, that encode transcription factors. Specifically, it
CC refers to transcription factor proteins that are capable of binding to
CC DNA in order to regulate gene transcription and gene expression in a
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC invention describes DNA constructs containing DNA encoding a
CC transcription factor that regulates the promoter, which is operably
CC linked to the desired nucleic acid to be expressed. It further provides
CC transgenic plants expressing a transcription factor that confers a trait
CC to the plant such as increased drought, salt or disease tolerance, height
CC change, enhanced cold/ frost tolerance, enhanced color, health and
CC nutritional characteristics, as well as improved taste, starch
CC composition, flower longevity and germination, amongst others.
CC Accordingly, such plants that are successfully transfected with a DNA
CC construct can be characterized by a difference in flower color, petal or
CC leaf shape and size, aroma or plant height. This polypeptide is a plant
CC transcription factor protein sequence of the invention.
XX
SQ Sequence 418 AA;
Query Match 53.5%; Score 1049.5; DB 9; Length 418;
Best Local Similarity 54.8%; Pred. No. 2.9e-108; Indels 31; Gaps 9;
Matches 216; Conservative 57; Mismatches 90;
QY 11 RSRPHRVHAAASTANSSD-----PFSWSELPEELLREILIRVETVGGDWPSRR 60
DB 32 RSRSHS--HGVAVTSGDENHCKTAADVLSHCWANLPYELLARDIIERIESSE-IWPAARR 88
QY 61 NVVACAGVCRWRILTKETIVAVPEFSSKLTFFPISLKQSGRDSLVOCFTKRNNTQSYHL 120
DB 89 NVVACAGVCRWRILTKETIVAVPEFSSKLTFFPISLKQSGRDSLVOCFTKRNNTQSYHL 148
QY 121 YGLTSLTDNGKFLAASKLRATCTDYIISLRSDISKRSNAYLGRMSNPLGTGKFTV 180
DB 149 YGLTSLTDNGKFLAASKLRATCTDYIISLRSDISKRSNAYLGRMSNPLGTGKFTV 208
QY 181 FPGSOT---GAAMQKRSRNNFKVSPRVPQSGYPIAHISYELNVLGSGRPMRCIMDT 237
DB 209 YDSQSHSGVATSHAGRVGSKQVSPRPVASSYVVAHAYELNVLGTRGPRMQCTMHS 268
QY 238 IPMSIVE-----SRGVASTSISFSS--RSSEVFRSHSKPLRSNASCSDSGNNIGDPP 290
DB 269 IPASVVEGGLNTSTPEIPGSTLVEPLASFRSS---RSKSVISESNAFSGPLNSGKLDKP 325
QY 291 LVLSNKAPRWHQOLRCWCLNFHGRVTVASVKQFQVAVSDCEAGQTS-----RIILOFGK 346
DB 326 LILKNKAPRWHEQLOQCWCLNFQGRVTVASVKQFQVAVSDCEAGQTS-----RIILOFGK 384
QY 347 VCKDMFTMDYGPISAFQAFALCLSSFEFRIACE 380
DB 385 VCKDMFTMDYGPISAFQAFALCLSSFEFRIACE 418
RESULT 15
ADW18506
ID ADW18506 standard; protein; 414 AA.
XX
AC ADW18506;
XX
DT 24-MAR-2005 (first entry)
DE Pinus radiata transcription factor protein TUBBY family Seq 2296.
XX
KW plant; transcription; gene regulation; gene expression; transgenic plant;
KW drought resistance; disease resistance; salt tolerance; cold tolerance;
KW freezing tolerance; flowering; flavor enhancer; flower color.
XX
OS Pinus radiata.
XX

PN WO2005001050-A2.
XX
PD 06-JAN-2005.
XX
PF 07-JUN-2004; 2004WO-USO17965.
XX
PR 06-JUN-2003; 2003US-0476189P.
XX
PA (ARBO-) ARBORGEN LLC.
XX
PI Bloksberg LN, Bryant C, Connatt MB, Emerson SJ, Frost MJ;
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX WPI; 2005-075542/08.
DR N-PSDB; ADW18185.
XX
XX New polynucleotides isolated from plants encoding transcription factors,
PT and polypeptides encoded by such polynucleotides, useful for regulating
PT gene transcription and gene expression.
XX
PS Claim 31; SEQ ID NO 2296; 1265pp; English.
XX
CC This invention relates to novel isolated plant nucleic acid molecules, or
CC variants thereof, that encode transcription factors. Specifically, it
CC refers to transcription factor proteins that are capable of binding to
CC DNA in order to regulate gene transcription and gene expression in a
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC invention describes DNA constructs containing DNA encoding a
CC transcription factor that regulates the promoter, which is operably
CC linked to the desired nucleic acid to be expressed. It further provides
CC transgenic plants expressing a transcription factor that confers a trait
CC to the plant such as increased drought, salt or disease tolerance, height
CC change, enhanced cold/ frost tolerance, enhanced color, health and
CC nutritional characteristics, as well as improved taste, starch
CC composition, flower longevity and germination, amongst others.
CC Accordingly, such plants that are successfully transfected with a DNA
CC construct can be characterized by a difference in flower color, petal or
CC leaf shape and size, aroma or plant height. This polypeptide is a plant
CC transcription factor protein sequence of the invention.
XX
SQ Sequence 414 AA;
Query Match 52.4%; Score 1028; DB 9; Length 414;
Best Local Similarity 53.0%; Pred. No. 7.4e-106;
Matches 220; Conservative 57; Mismatches 102; Indels 36; Gaps 10;
QY 1 MTFRSLLOEM-----RSRPHRVV--HAAASTANGSDPPS-----WSEL 36
DB 1 MSLRSIMREIKVRNGIGNISRRRSDMRVSNHRRASQATDDTLRSLVLFVQOSQWANL 60
QY 37 PPELLREILIRVETVGGDWPSRRNVVACAGVCRWRILTKETIVAVPEFSSKLTFFPISLK 96
DB 61 PPELLHNVTQRLSEASE-GTWPARKNVVACAGVCKSWREITKEVKTPEQCKLTPFISLK 119
QY 97 QSGPRDSLVCQCFIKRNNTQSYHLVGLTTS--LTNGKFLAASKLRATCTDYIISLRS 155
DB 120 QSGPRDPPMQCFIKRDRSASTYKLFGLTTPALLVENGKFLFAKKVRRHATDYIISLS 179
QY 156 DDISKRSNAYLGRMSNPLGTGKFTVFDGS--OTGAAMQKRSRNNFI--KVSPRPVQGSY 211
DB 180 EDFSRASNTYVCKRNSNPLGTFTTYDSQPPHSGAVLTSRFRSRRFYKQVSPRPVAGSY 239
QY 212 PIAHISYELNVLGSGRPMRCIMDTIPMSIVESRGVAVASTSISFSSRSSSPVFRSHSKP 271
DB 240 NIANIVYELNVLGTRGPRMQCTMHSIPASSIQVGNAPTTEFFPRSLDESFLSPSKEP 299
QY 272 LRS-NSASCSDSGNNL--GDPLVLNKAAPRWHEQOLRCWCLNFHGRVTVASVKNFOLVA- 327
DB 300 LIDFSSSSLSLEIPINVESKDVPLVKNKAPRWHEQOLQCWCLNFHGRVTVASVKNFOLVAS 359
QY 328 --VSDCEAGQTSERILQFGKVDKMTMDYGYPTSAFOAFALCLSSFEFRIACE 380
XX

Db 360 VEPSQLVTQADQDKVILQFGKIGKDIFTMDYRYPPLSAFQAFALCLSSFDTKLACE 414

Search completed: December 23, 2005, 23:23:52
Job time : 202 secs